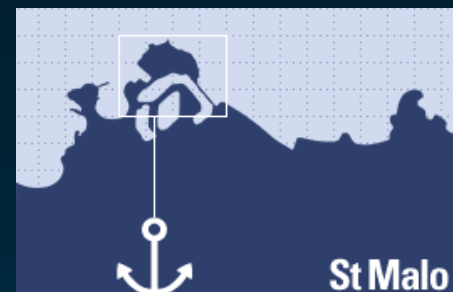Medical Informatics Europe 2003
St-Malo, France
May 4, 2003 - Tutorial T4

St Malo

The Unified Medical Language System
What is it and how to use it?

Olivier Bodenreider

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA

NATIONAL LIBRARY OF MEDICINE

Part I

What is the UMLS?

Outline

◆ Part I

- Introduction

- Overview through an example

- UMLS Metathesaurus

- UMLS Semantic Network

- SPECIALIST lexicon and lexical tools

Introduction

Motivation

[Lindberg & al., *Methods*, 1993]
[Humphreys & al., *JAMIA*, 1998]

- Started in 1986

- National Library of Medicine

- "Long-term R&D project"

- Complementary to IAIMS   (Integrated Academic Information Management Systems)

«[…] the UMLS project is an effort to overcome two significant barriers to effective retrieval of machine-readable information.
- The first is the variety of ways the same concepts are expressed in different machine-readable sources and by different people.
- The second is the distribution of useful information among many disparate databases and systems.»

NLM

5

UMLS chronology

- Definition of 3 knowledge sources (1986-88)
  - Metathesaurus
  - Semantic Network
  - Information Sources Map
- Building, distributing, and testing (1989-91)
  - Integration vs. *ad hoc* development
  - First release in 1990
- Development of applications (1992-94)

Terminology Adrenal gland diseases
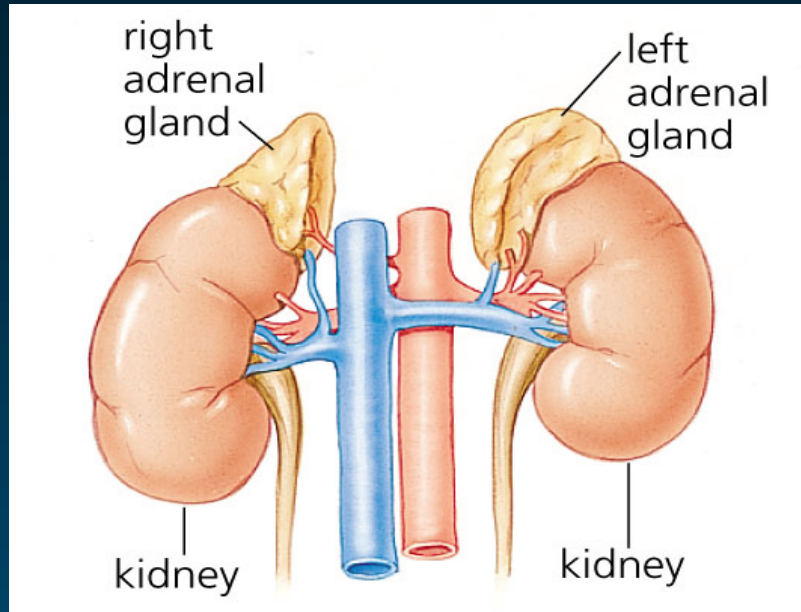
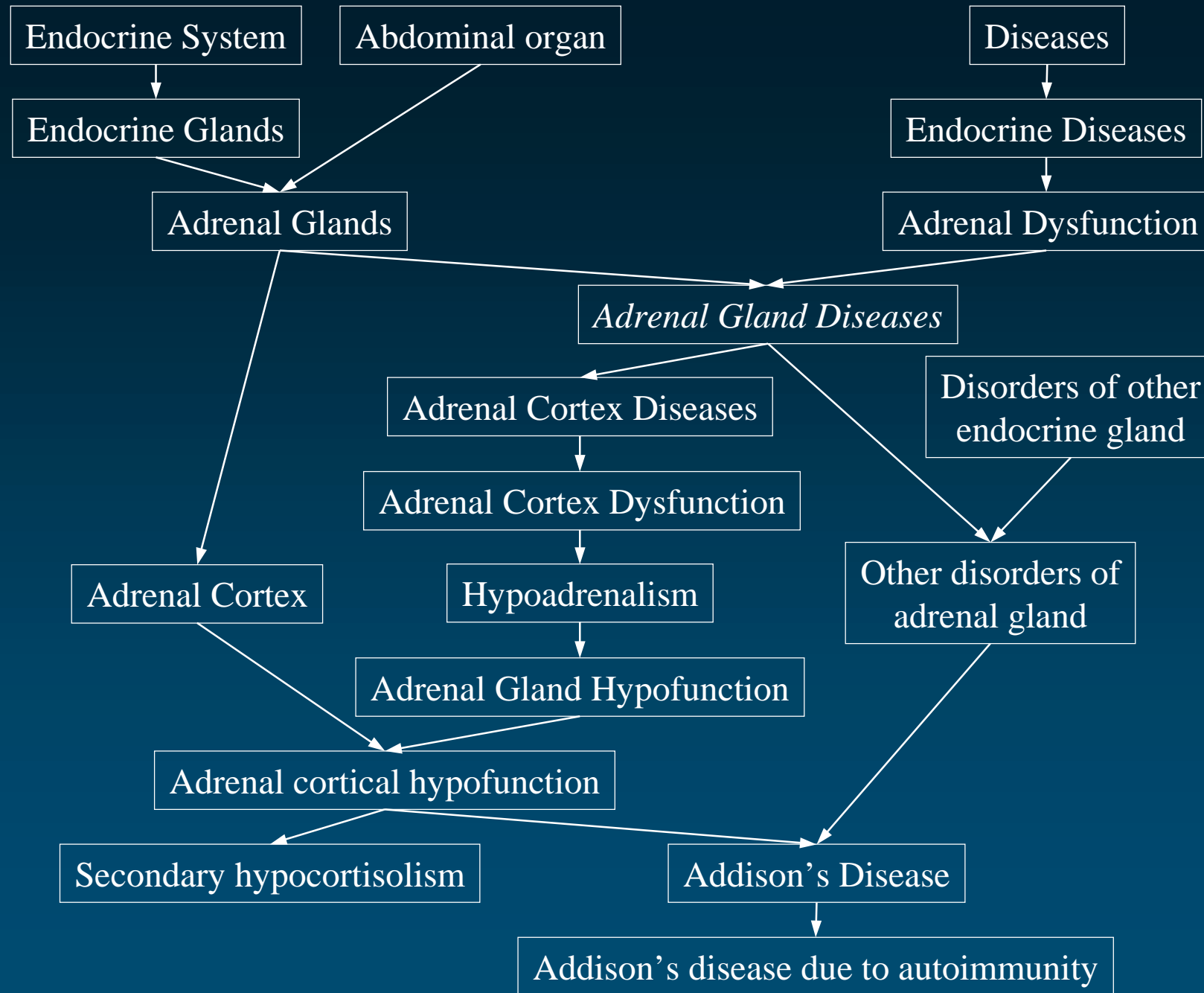|		
Adrenal gland diseases	MeSH	D000307
Adrenal disorder	AOD	0000005418
Disorder of adrenal gland	Read	C15z.
Diseases of the adrenal glands	SNOMED	DB-70000

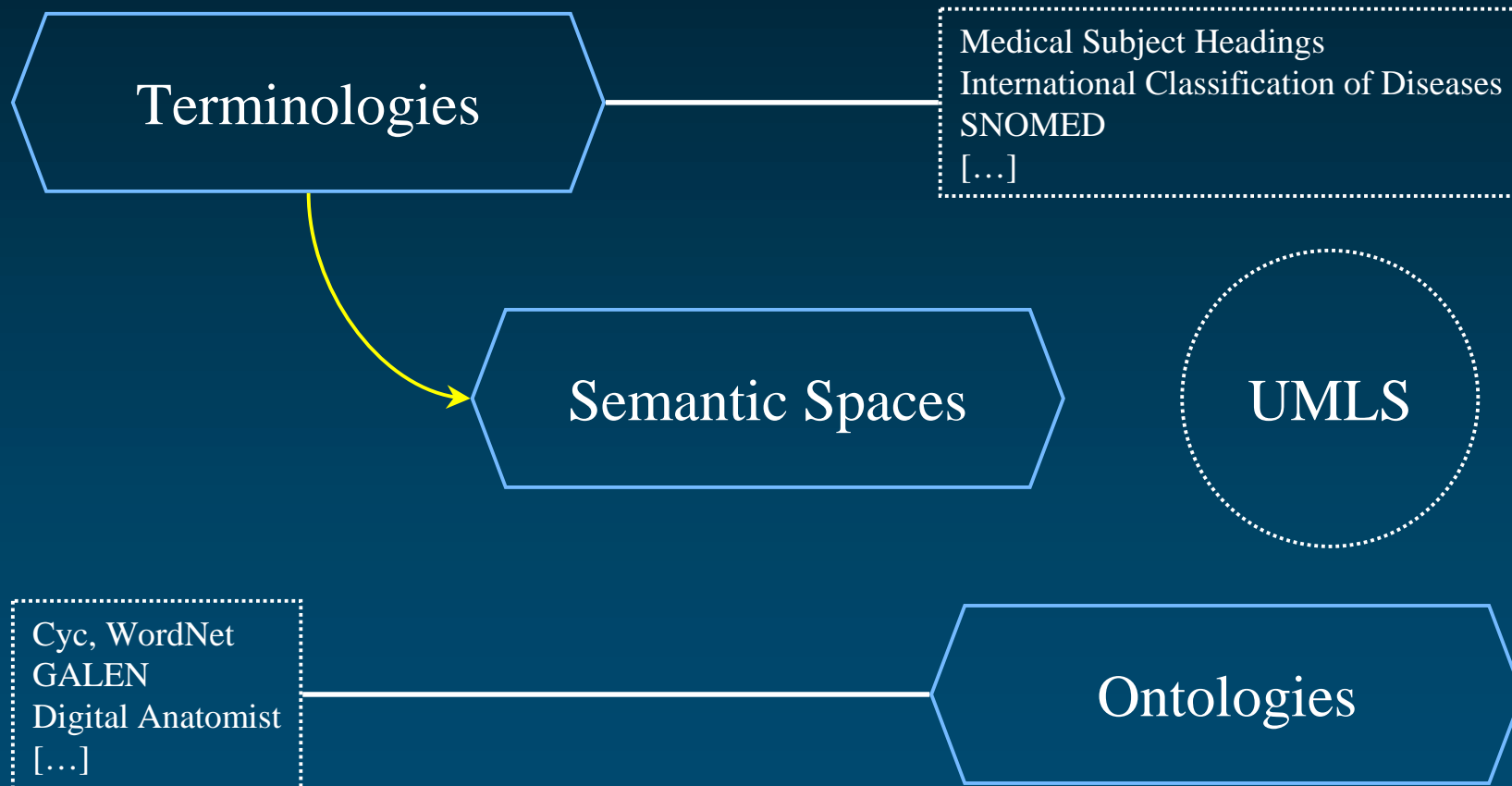# UMLS Adrenal gland diseases concept

Disease or
Syndrome

Adrenal Gland Diseases

C0001621

Adrenal gland diseases	MeSH	D000307
Adrenal disorder	AOD	0000005418
Disorder of adrenal gland	Read	C15z.
Diseases of the adrenal glands	SNOMED	DB-70000

NLM

8

Biomedical knowledge organization

Terminologies

Medical Subject Headings
International Classification of Diseases
SNOMED
[…]

Semantic Spaces

UMLS

Cyc, WordNet
GALEN
Digital Anatomist
[…]

Ontologies
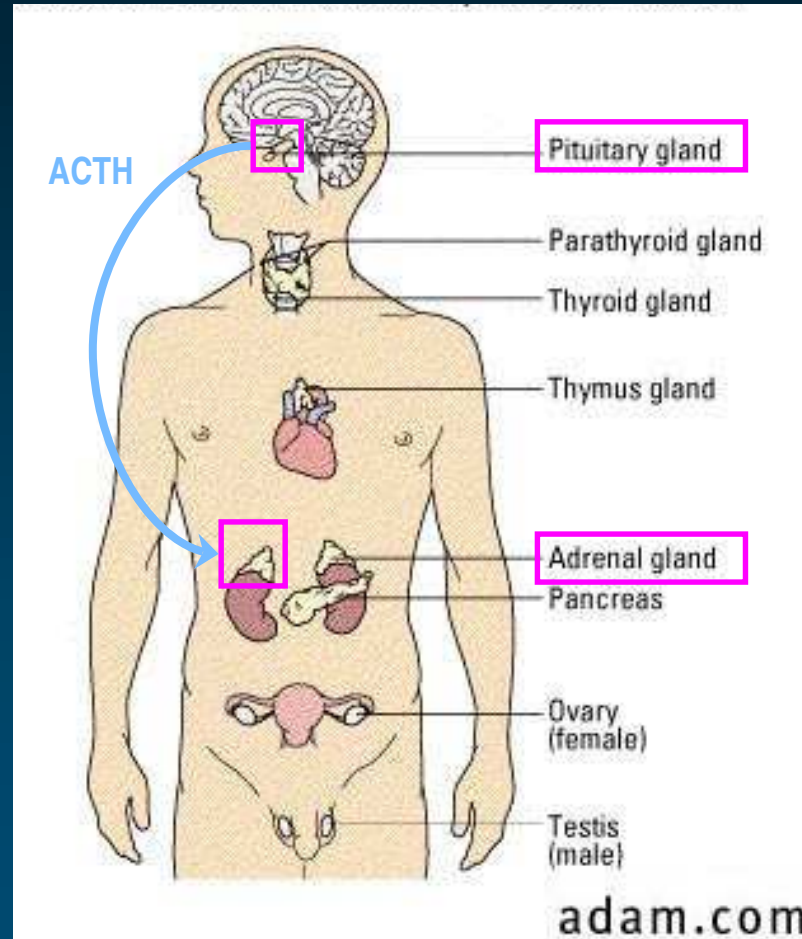
NLM

10

Overview through an example

Addison's disease

◆ Addison's disease is a rare endocrine disorder

◆ Addison's disease occurs when the adrenal glands do not produce enough of the hormone cortisol

◆ For this reason, the disease is sometimes called chronic adrenal insufficiency, or hypocortisolism

cortex   medulla

right adrenal gland

left adrenal gland

kidney   kidney
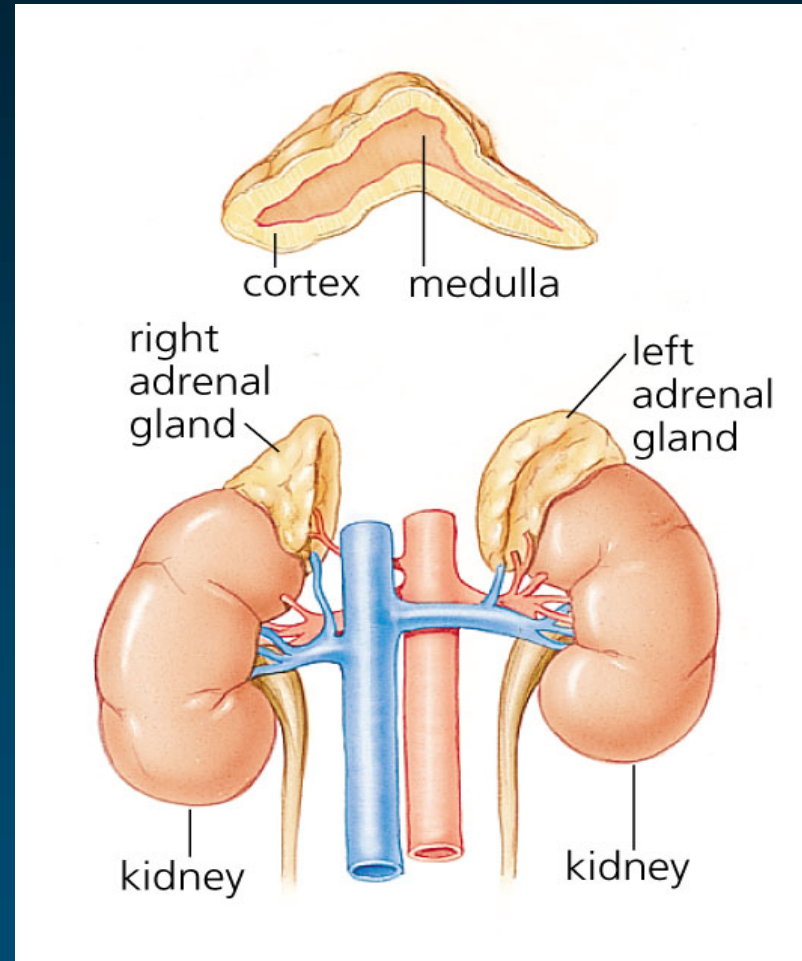
Adrenal insufficiency Clinical variants

- ◆ Primary / Secondary
  - ● Primary: lesion of the adrenal glands themselves
  - ● Secondary: inadequate secretion of ACTH by the pituitary gland
- ◆ Acute / Chronic
- ◆ Isolated / Polyendocrine deficiency syndrome

ACTH

Pituitary gland

Parathyroid gland

Thyroid gland

Thymus gland

Adrenal gland

Pancreas

Ovary (female)

Testis (male)

adam.com

Addison's disease: Symptoms

- Fatigue

- Weakness

- Low blood pressure

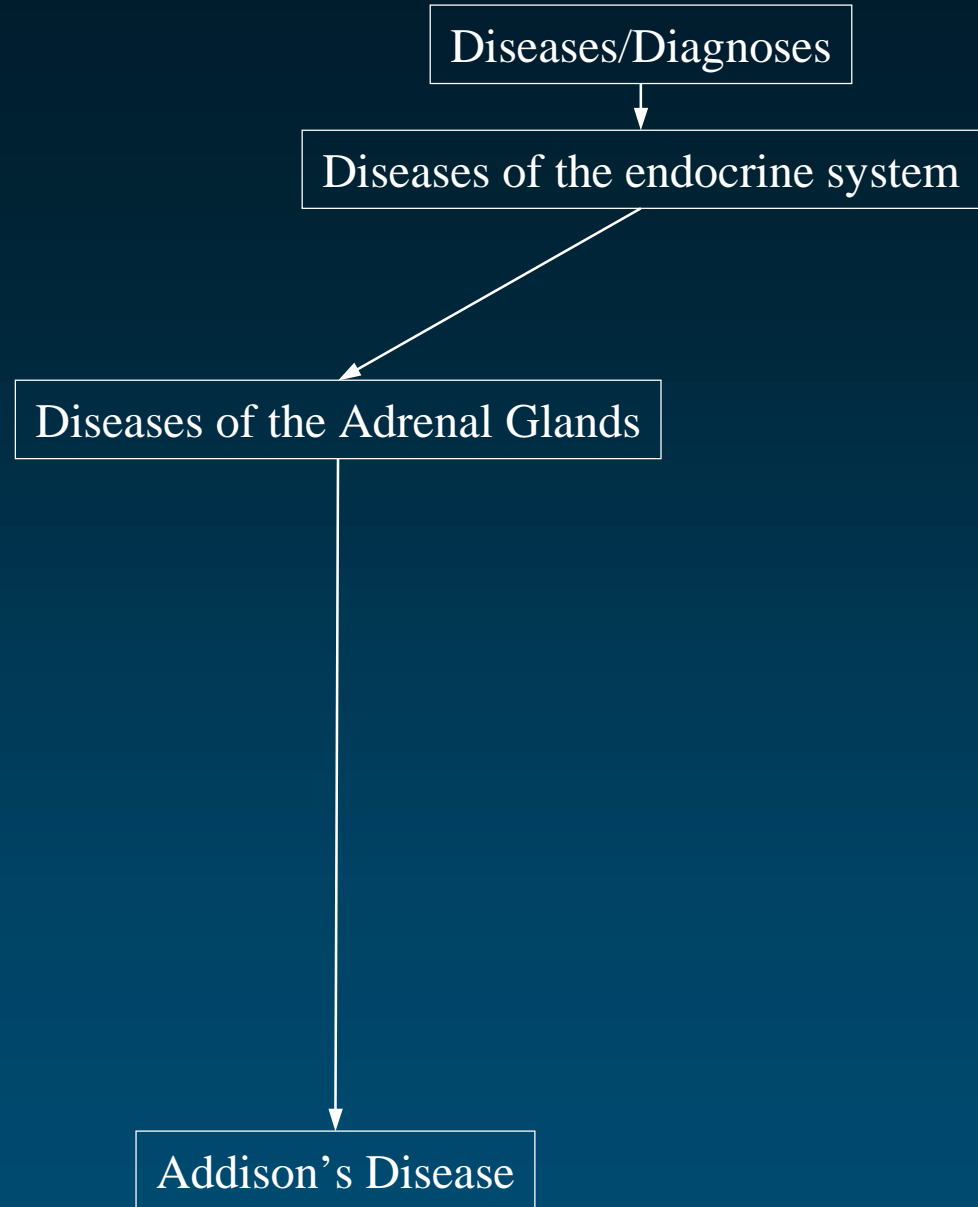- Pigmentation of the skin (exposed and non-exposed parts of the body)

- …

AD in medical vocabularies

◆ Synonyms: different terms

- Addisonian syndrome   ] eponym
- Bronzed disease
- Addison melanoderma   symptoms
- Asthenia pigmentosa
- Primary adrenal deficiency
- Primary adrenal insufficiency
- Primary adrenocortical insufficiency   clinical variants
- Chronic adrenocortical insufficiency

◆ Contexts: different hierarchies

MeSH

Diseases

Endocrine Diseases

Adrenal Gland Diseases

Adrenal Gland Hypofunction

Addison's Disease

Endocrine disorder

Adrenal disorder

Adrenal cortical disorder

Adrenal cortical hypofunction

Addison's Disease

Read Codes

Endocrine disorder

Disorder of adrenal gland

Hypoadrenalism

Adrenal Hypofunction

Corticoadrenal insufficiency

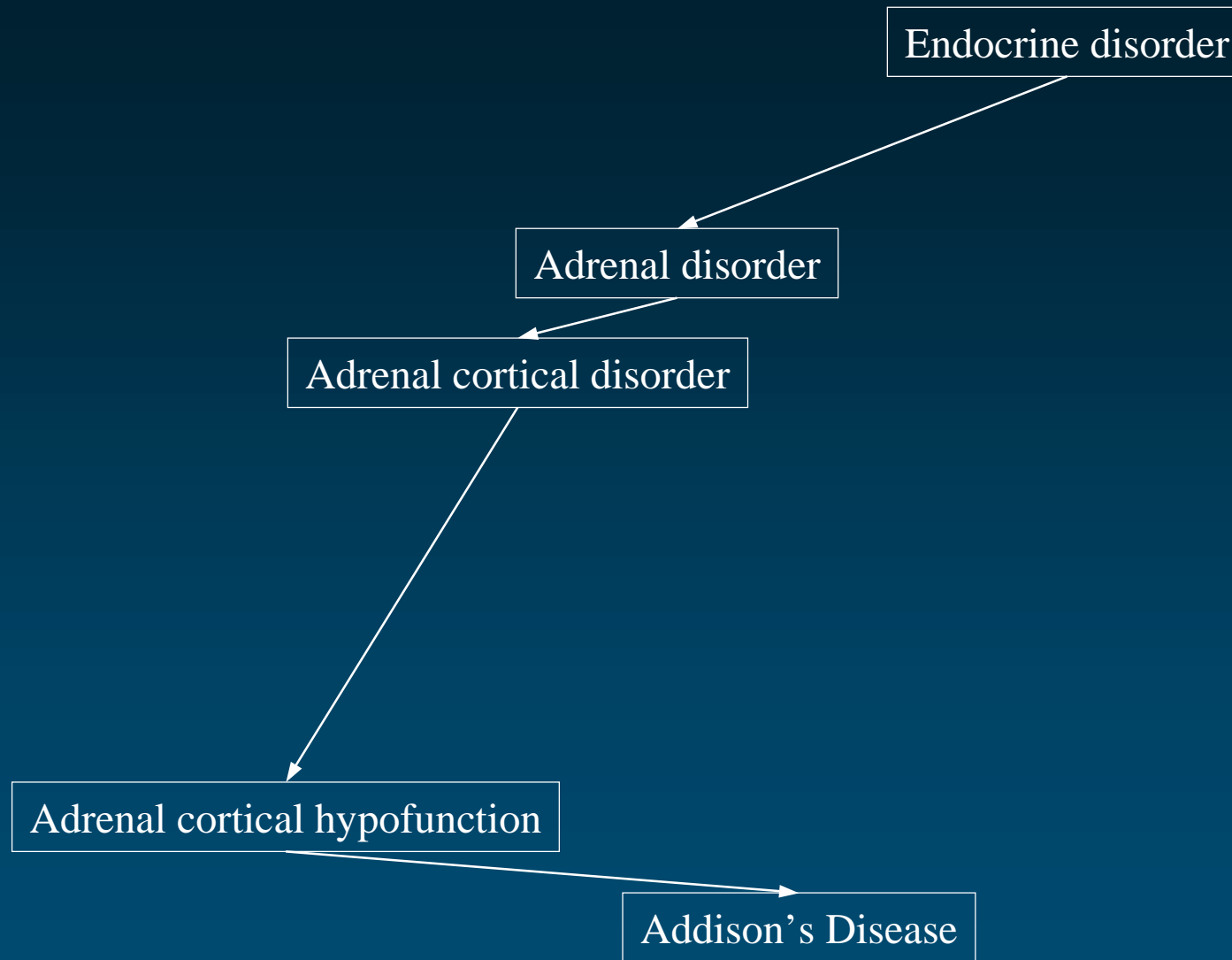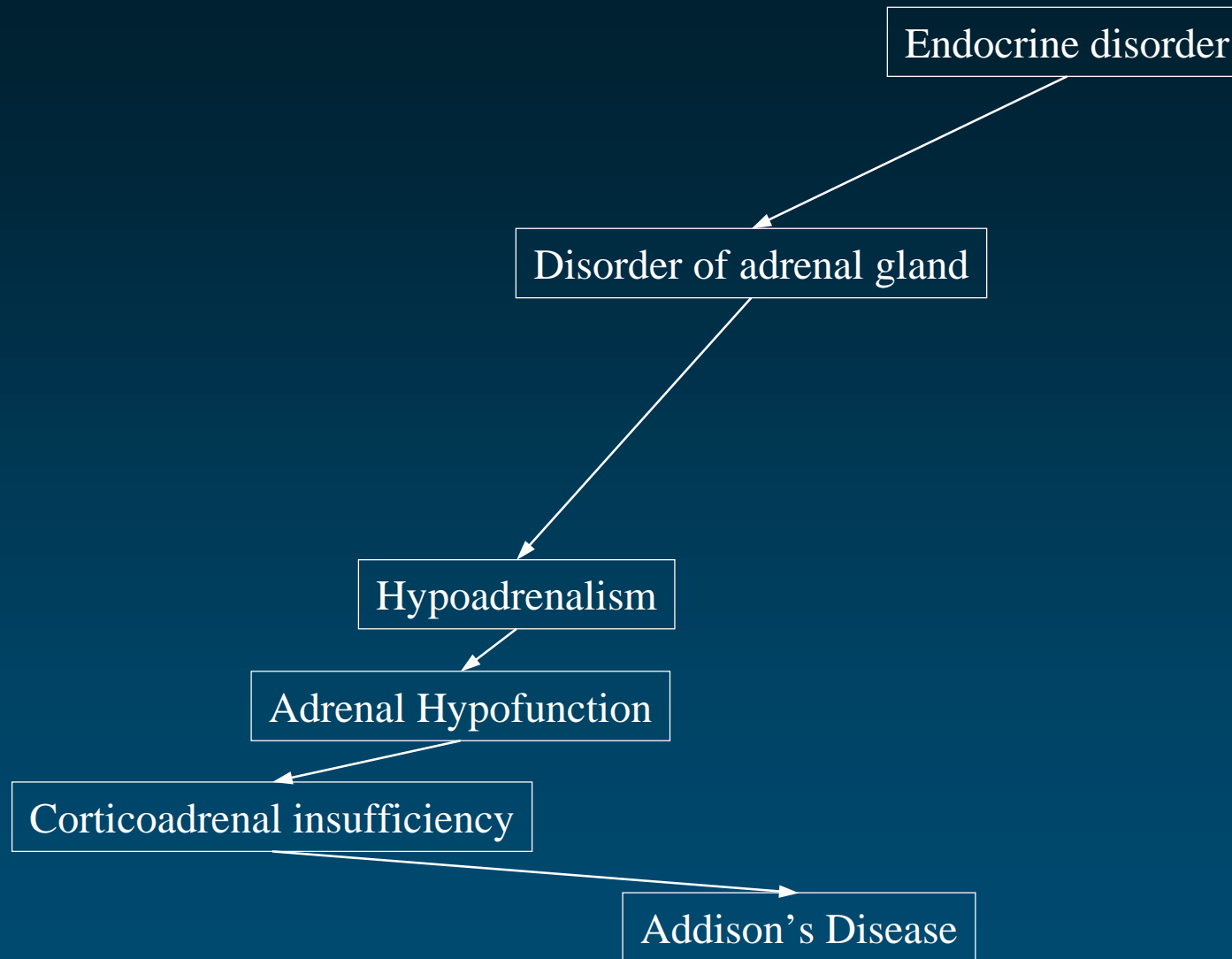Addison's Disease
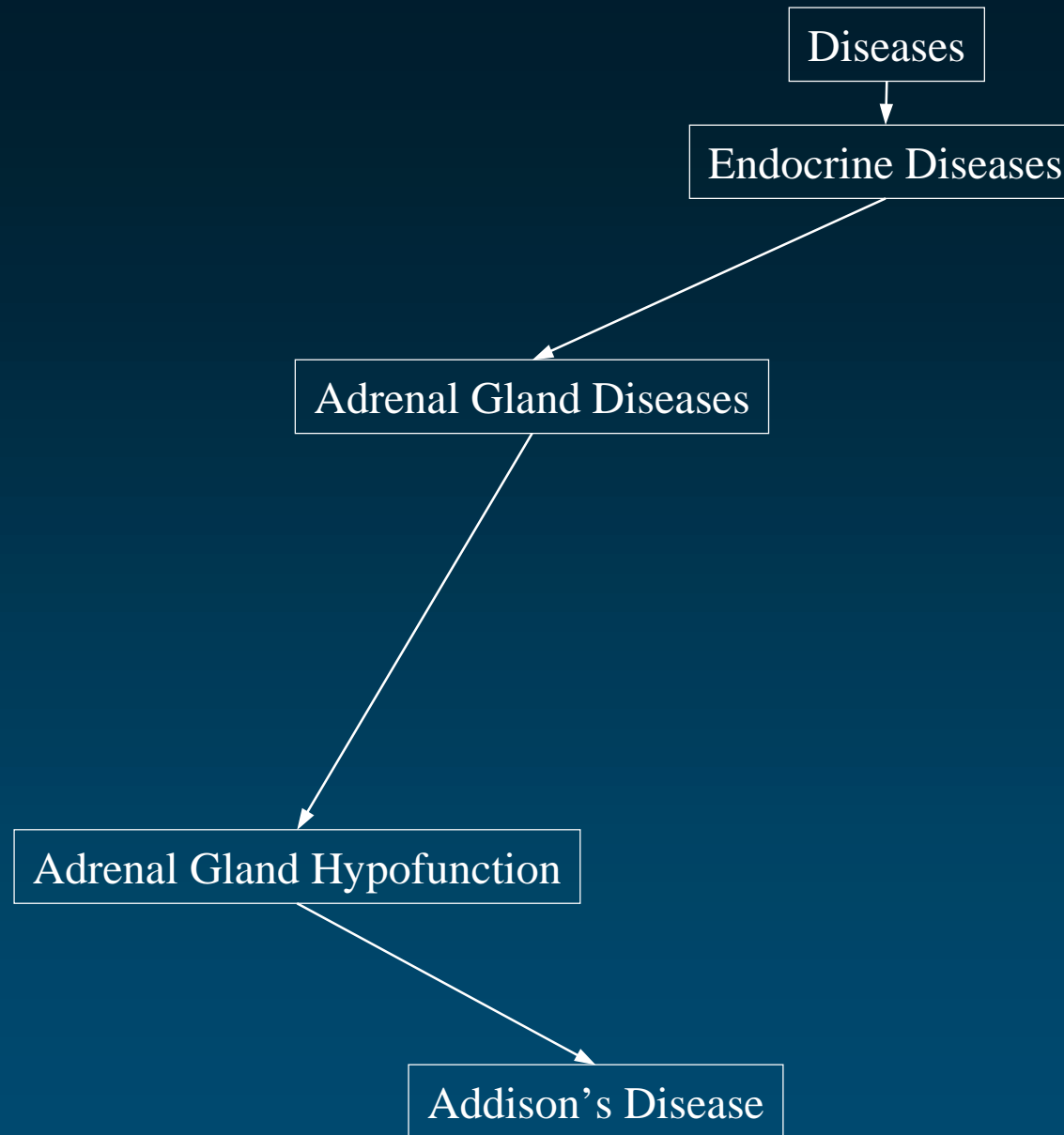
ICD-10

Disorders of other endocrine gland

Other disorders of adrenal gland

Primary adrenocortical insufficiency

From the vocabularies to the UMLS

- Vocabularies provide
  - terms
  - hierarchies

- Organize terms

- Organize concepts

- Relate to other concepts

- Metathesaurus = Thesaurus of Thesauri

Organize terms

◆ Synonymous terms clustered into a concept

◆ Preferred term

◆ Unique identifier (CUI)

Adrenal gland diseases	MeSH	D000307
Adrenal disorder	AOD	0000005418
Disorder of adrenal gland	Read	C15z.
Diseases of the adrenal glands	SNOMED	DB-70000

C0001621

Adrenal Gland Diseases

organize terms

Diseases

Endocrine diseases
Endocrine disorder
C0014130

Endocrine Diseases

Adrenal gland diseases
Adrenal disorder
Disorder of adrenal gland
Diseases of the adrenal glands
C0001621

Adrenal Gland Diseases

Disorders of other endocrine gland

Adrenal Cortex Diseases

Adrenal cortex diseases
Adrenal cortical disorder
C0001614

Hypoadrenalism

Other disorders of adrenal gland

Adrenal gland hypofunction
Adrenal hypofunction
C0001623

Adrenal Gland Hypofunction

Adrenal cortical hypofunction

Adrenal cortical hypofunction
Adrenocortical insufficiency
C0405580

Addison's disease
Primary adrenocortical insufficiency
C0001403

Addison's Disease

Organize concepts

- Inter-concept relationships: hierarchies from the source vocabularies

- Redundancy: multiple paths

- One graph instead of multiple trees (multiple inheritance)

organize concepts

Endocrine Diseases

Adrenal Gland Diseases

Adrenal Cortex Diseases

SNOMED

MeSH

AOD

Read Codes

Hypoadrenalism

Adrenal Gland Hypofunction

Adrenal cortical hypofunction

Addison's Disease

organize concepts

Endocrine Diseases

Adrenal Gland Diseases

Adrenal Cortex Diseases

SNOMED
MeSH
AOD
Read Codes

Hypoadrenalism

Adrenal Gland Hypofunction

UMLS

Adrenal cortical hypofunction

Addison's Disease
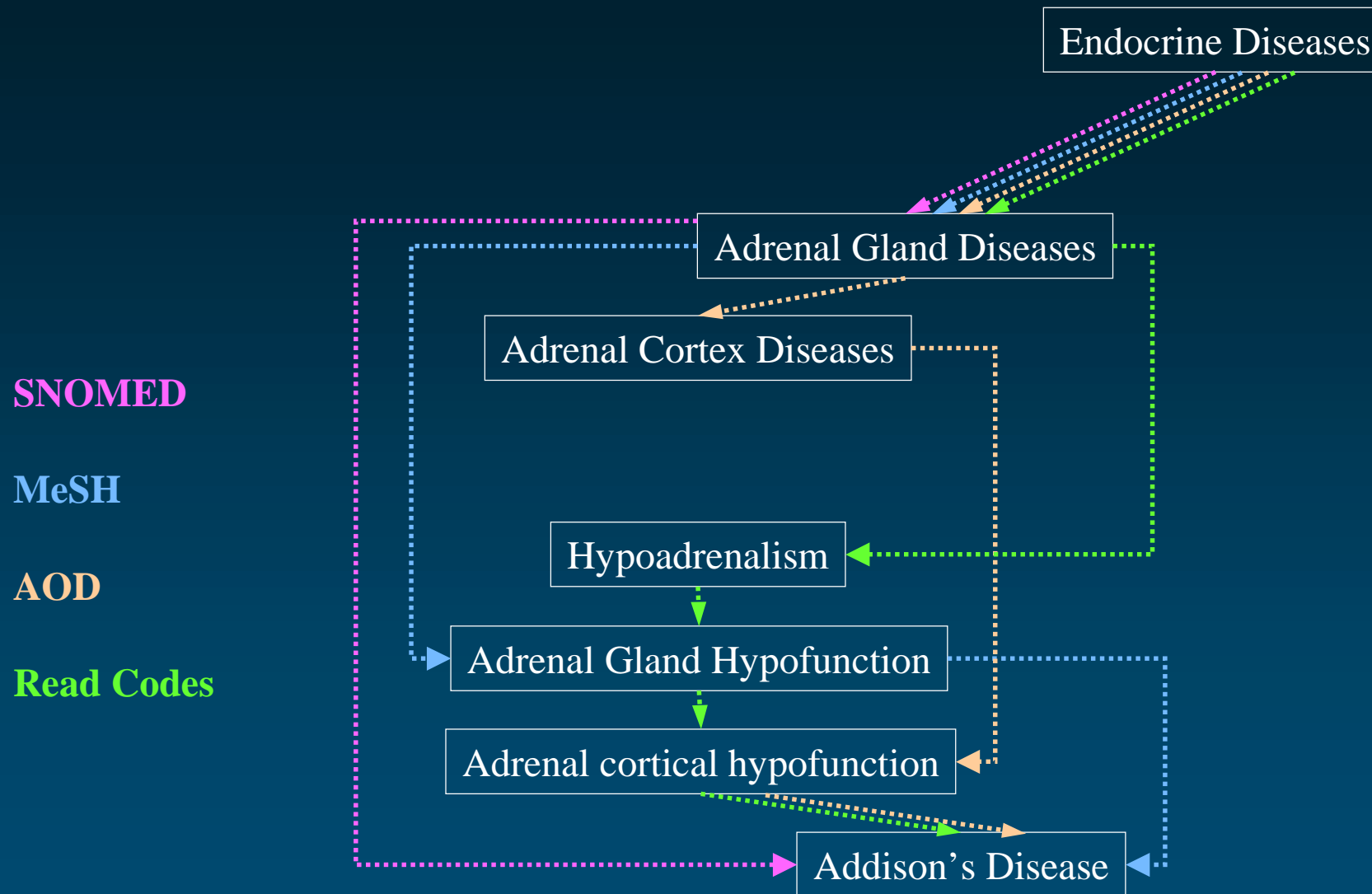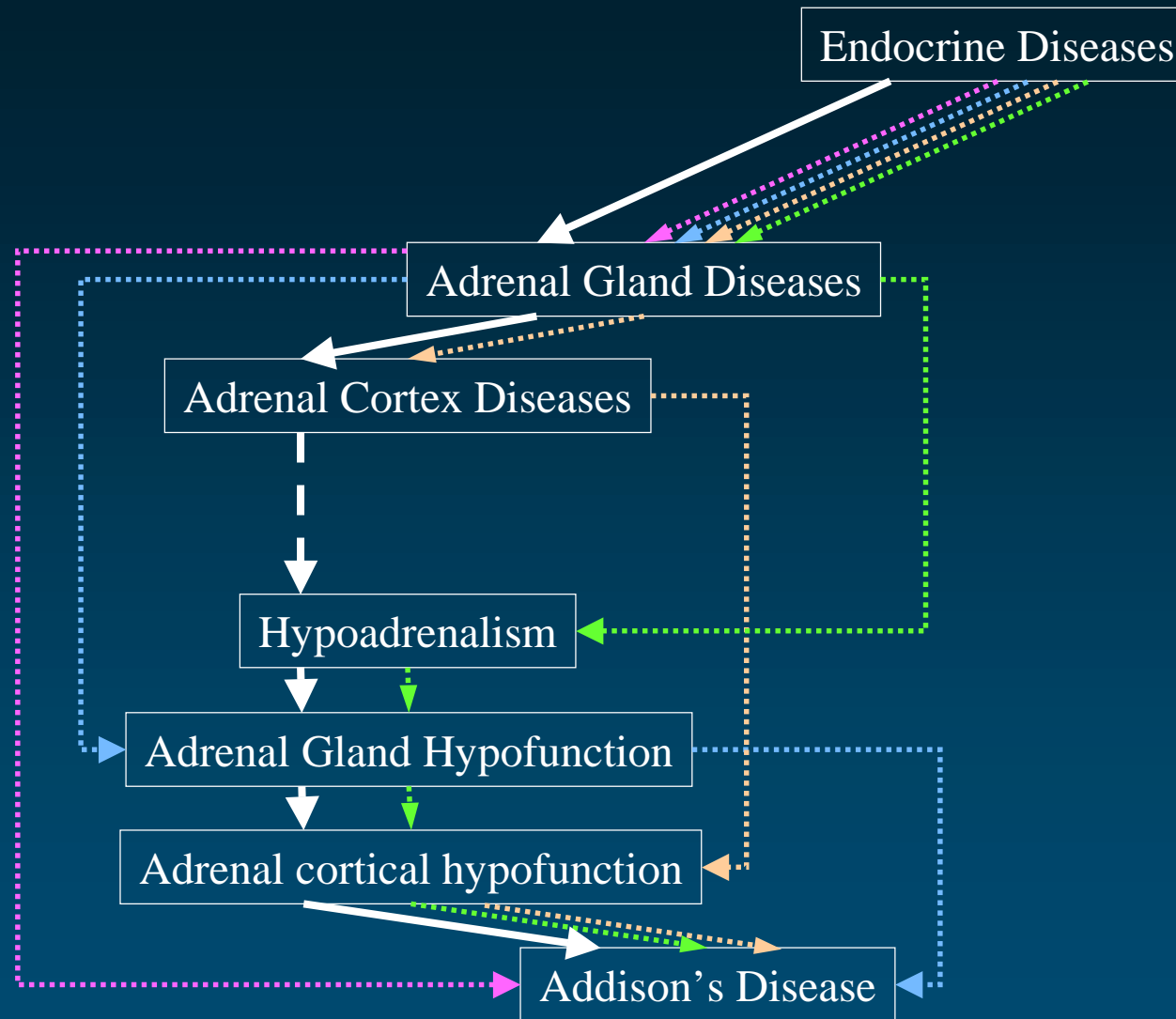
Relate to other concepts

- Additional hierarchical relationships
  - link to other trees
  - make relationships explicit
- Non-hierarchical relationships
- Co-occurring concepts
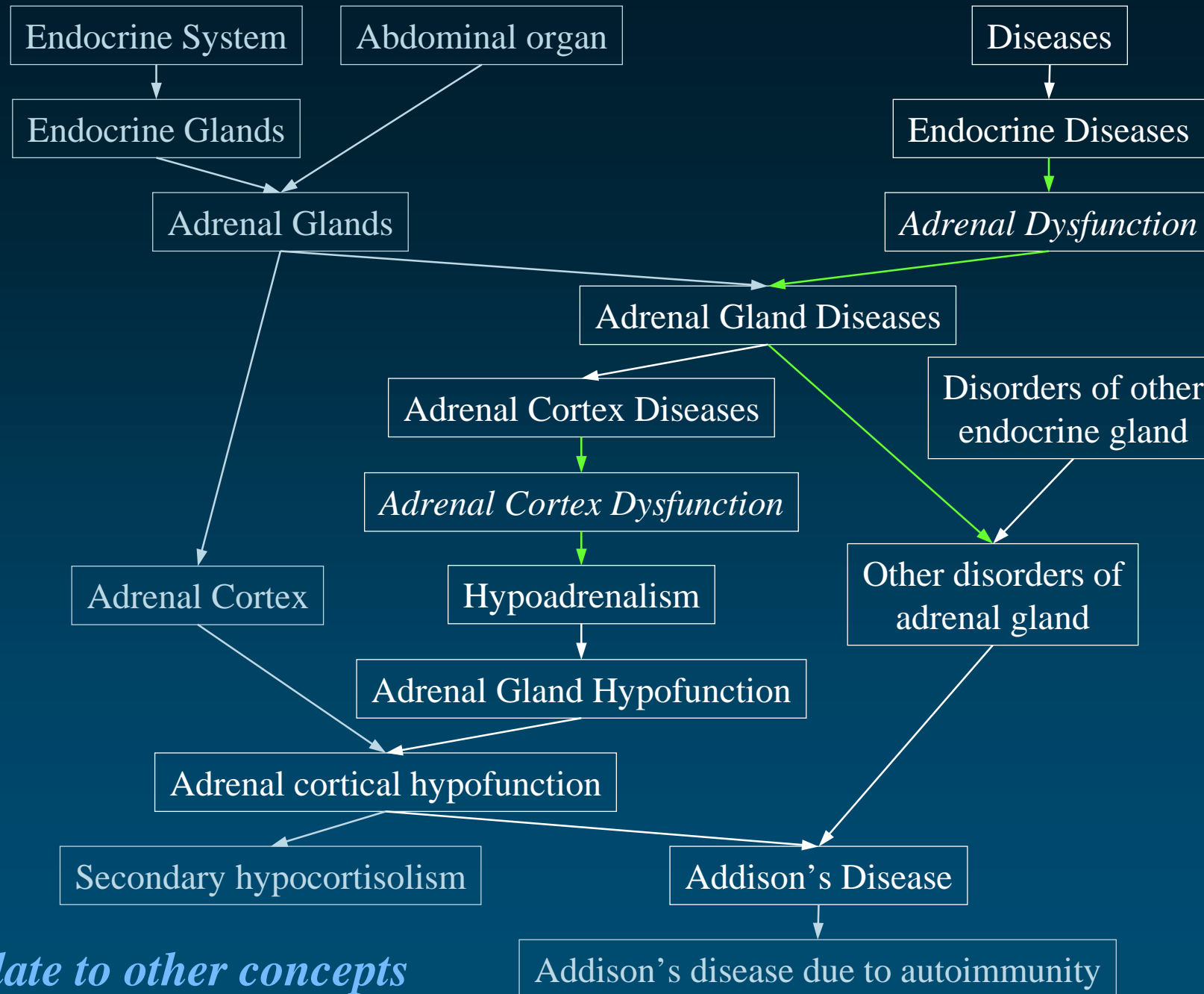
Higher-level organization

◆ Semantic types: broad categories
- Disease or Syndrome
- Body Part, Organ, or Organ Component

◆ Semantic relationships
- hierarchical: is a kind of (isa)
- non-hierarchical (location_of, caused_by)

◆ Semantic network (SN = STs + SRs)

◆ Semantic categorization
- each concept is given (at least) one ST

Semantic Network

Semantic Types

Fully Formed Anatomical Structure

location of

Biologic Function

isa

Pathologic Function

isa

isa

Body Part, Organ or Organ Component

Disease or Syndrome

isa

Concepts

Adrenal Glands

Adrenal Cortex

Adrenal cortical hypofunction

Addison's Disease

Metathesaurus

How do they do that?

- Lexical knowledge
- Lexical resources
  - Lexicon
  - Lexical programs
- UMLS editors

Lexical knowledge

Adrenal gland diseases
Adrenal disorder
Disorder of adrenal gland
Diseases of the adrenal glands
C0001621

Diseases

Endocrine Diseases

Adrenal cortical hypofunction

Adrenal cortical hypofunction
Adrenocortical insufficiency
C0405580

Addison's disease
Primary adrenocortical insufficiency
C0001403

Addison's Disease
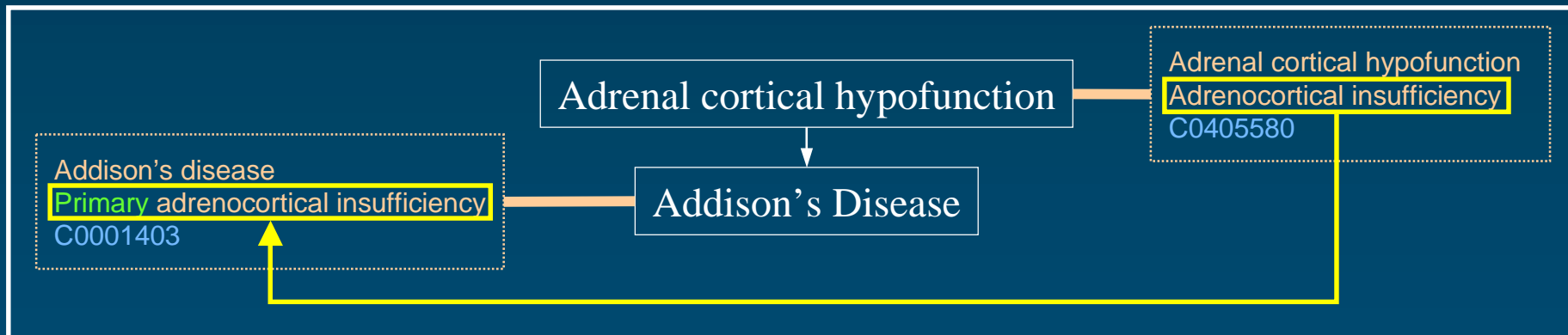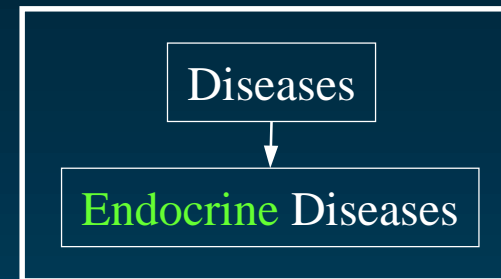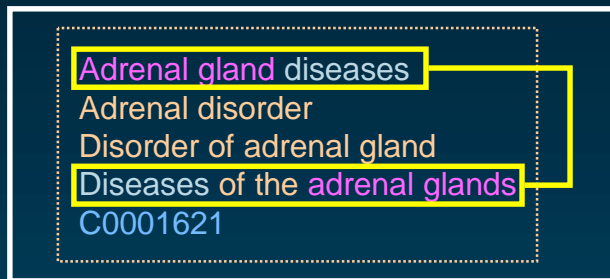
Lexical resources

◆ Lexicon

Syntactic Category: noun
Inflection Type: reg

Base Form: gland

Singular: gland
Plural: glands

◆ Lexical tools

● stop words   Diseases of the adrenal glands

● word order   Diseases of the adrenal glands
                              Adrenal glands diseases

● inflection   gland ⟶ glands

● derivation   cortex ⟶ cortical

Additional knowledge: UMLS editors
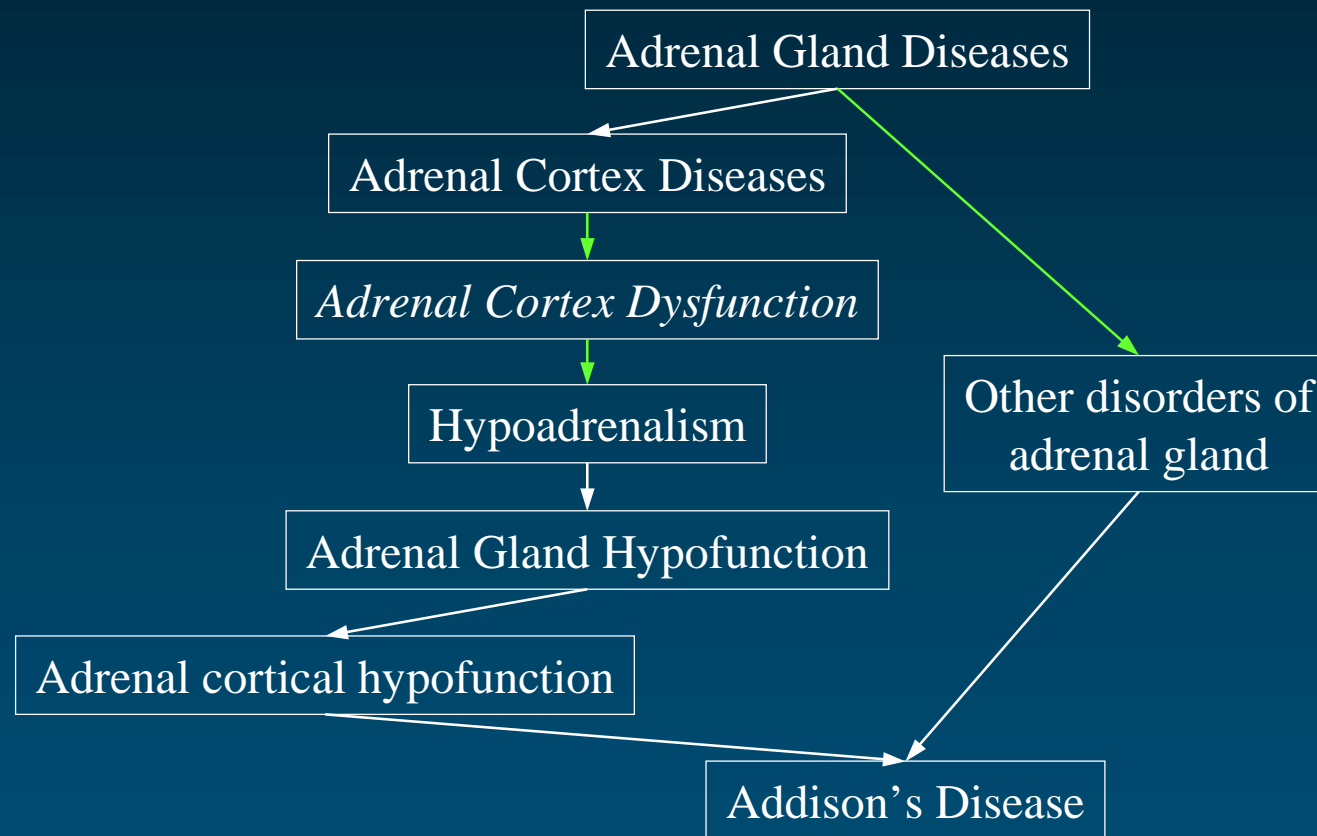
AD in the UMLS

- Synonymous terms clustered into concepts
- Unique identifier
- Finer granularity
- Broader scope
- Additional hierarchical relationships
- Semantic categorization

UMLS knowledge sources

UMLS: 3 components

- ◆ Metathesaurus
  - Concepts
  - Inter-concept relationships
- ◆ Semantic Network
  - Semantic types
  - Semantic network relationships
- ◆ Lexical resources
  - SPECIALIST Lexicon
  - Lexical tools

37

UMLS Metathesaurus

Metathesaurus Basic organization

- Terms / Concepts
  - Synonymous terms are clustered into a concept
  - Properties are attached to concepts, e.g.,
    - Unique identifier
    - Definition
- Relations
  - Concepts are related to other concepts
  - Properties are attached to relations, e.g.,
    - Type of relationship
    - Source

Source Vocabularies

- 117 "sources"
- ~60 families of vocabularies
  - multiple translations (e.g., MeSH, ICPC, ICD-10)
  - variants (American-English equivalents, Australian extension/adaptation)
  - subsequent versions usually considered distinct families (ICD: 9-10; DSM: IIIR-IV)
- Broad coverage of biomedicine
- Common presentation

Biomedical terminologies

- Core vocabularies
  - anatomy (UWDA, Neuronames)
  - drugs (First DataBank, Micromedex)
  - medical devices (UMD, SPN)
- Several perspectives
  - clinical terms (SNOMED, CTV3)
  - information sciences (MeSH, CRISP)
  - administrative terminologies (ICD-9-CM, CPT-4)
  - standards (HL7, LOINC)

Biomedical terminologies (cont'd)

- Specialized vocabularies
  - nursing (NIC, NOC, NANDA, Omaha, PCDS)
  - dentistry (CDT)
  - oncology (PDQ)
  - psychiatry (DSM, APA)
  - adverse reactions (COSTART, WHO ART)
  - primary care (ICPC)
- Knowledge bases (AI/Rheum, DXplain, QMR)

NLM

Addison's Disease: Concept

Disease or Syndrome

Addison's Disease

ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)
ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE
Addison melanoderma
Melasma addisonii
Primary adrenal deficiency
Asthenia pigmentosa
Bronzed disease
Insufficiency, adrenal primary
Primary adrenocortical insufficiency
Addison's, disease

MALADIE D'ADDISON - French
Addison-Krankheit - German
Morbo di Addison - Italian
DOENCA DE ADDISON - Portuguese
ADDISONOVA BOLEZN' - Russian
ENFERMEDAD DE ADDISON - Spanish

SNOMED
MeSH
AOD
Read Codes
…

C0001403

A disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. In the absence of replacement therapy, it is usually fatal.

NLM

43

Metathesaurus Concepts
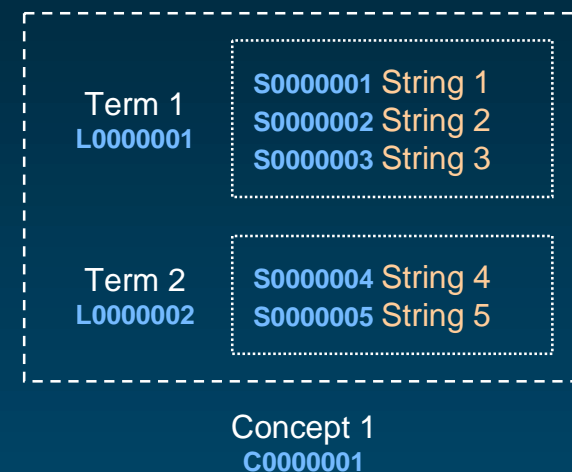
- ◆ Concept: Cluster of synonymous terms
  - ● ~875,000 concepts
  - ● identified by a CUI
- ◆ Term: Set of lexical variants
  - ● ~1.8 M terms
  - ● identified by a LUI
- ◆ String: Concept name
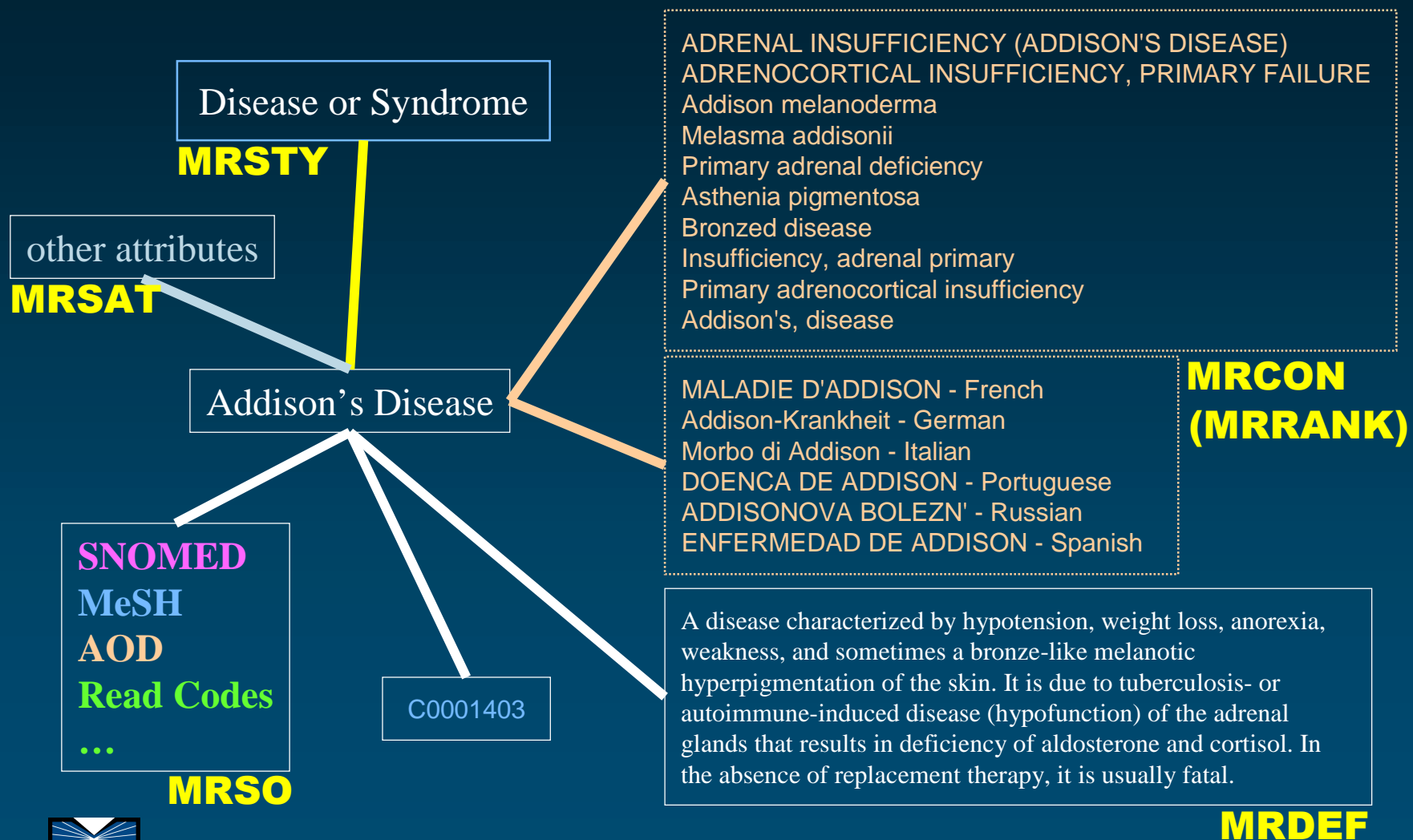 - ● ~2.1 M strings
  - ● identified by a SUI

Term 1
L0000001

S0000001 String 1
S0000002 String 2
S0000003 String 3

Term 2
L0000002

S0000004 String 4
S0000005 String 5

Concept 1
C0000001

Cluster of synonymous terms

Concept
C0001621

Term
L0001621

S0011232 *Adrenal Gland Diseases*
S0011231 Adrenal Gland Disease
S0000441 Disease of adrenal gland
S0481705 Disease of adrenal gland, NOS
S0220090 Disease, adrenal gland
S0044801 Gland Disease, Adrenal

[…]

Term
L0041793

S0860744 *Disorder of adrenal gland, unspecified*
S0217833 Unspecified disorder of adrenal glands

Term
L0161347

S0225481 *ADRENAL DISORDER*
S0627685 DISORDER ADRENAL (NOS)

[…]

Term
L0181041

S0632950 *Disorder of adrenal gland*
S0354509 Adrenal Gland Disorders

[…]

Term
L0368399

S0586222 *Adrenal disease*
S0466921 ADRENAL DISEASE, NOS

[…]

Term
L1279026

S1520972 *Nebennierenkrankheiten*

GER

Term
L0162317

S0226798 *SURRENALE, MALADIES*

FRE

[…]

NLM

Metathesaurus files Concepts

Disease or Syndrome

MRSTY

other attributes

MRSAT

Addison's Disease

SNOMED
MeSH
AOD
Read Codes
…

MRSO

C0001403

ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)
ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE
Addison melanoderma
Melasma addisonii
Primary adrenal deficiency
Asthenia pigmentosa
Bronzed disease
Insufficiency, adrenal primary
Primary adrenocortical insufficiency
Addison's, disease

**MRCON
(MRRANK)**

MALADIE D'ADDISON - French
Addison-Krankheit - German
Morbo di Addison - Italian
DOENCA DE ADDISON - Portuguese
ADDISONOVA BOLEZN' - Russian
ENFERMEDAD DE ADDISON - Spanish

A disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. In the absence of replacement therapy, it is usually fatal.

MRDEF

46

NLM

Metathesaurus Files

- Self-documentation
  - Files   MRFILES
  - Columns   MRCOLS
- Concept properties
  - Set of terms   MRCON
  - List of sources (+ original identifiers)   MRSO
  - Definition(s)   MRDEF
  - Semantic type(s)   MRSTY
  - Associated expression(s)   MRATX

Metathesaurus Files (continued)

- ◆ More concept properties
  - Contexts   MRCXT
  - String attributes   MRSAT
  - Locators   MRLO
  - Term ranking   MRRANK
- ◆ Indexes
  - Word indexes   MRXW.XXX
  - Normalized indexes   MRXNW.ENG   MRXNS.ENG

48

Metathesaurus Files (continued)

- ◆ Ambiguity   AMBIG.XUI
- ◆ Change files
  - ● Deleted   DELETED.XUI
  - ● Merged   MERGED.XUI
  - ● Retired   MRCUI
- ◆ Source information   MRSAB

Metathesaurus Evolution over time

- ◆ Concepts never die (in principle)
  - ● CUIs are permanent identifiers
- ◆ What happens when they do die (in reality)?
  - ● Concepts can merge or split
  - ● Resulting in new concepts and deletions

MRCUI

Addison's disease, NOS
C0271735

ADRENOCORTICAL INSUFFICIENCY,
PRIMARY FAILURE   C0241779

Addison's disease
C0001403

CUI1	VER	CREL	CUI2
C0241779	1996	SY	C0001403
C0271735	1996	SY	C0001403

1992   1993   1994   1995   1996   1997   1998   1999   …   2002
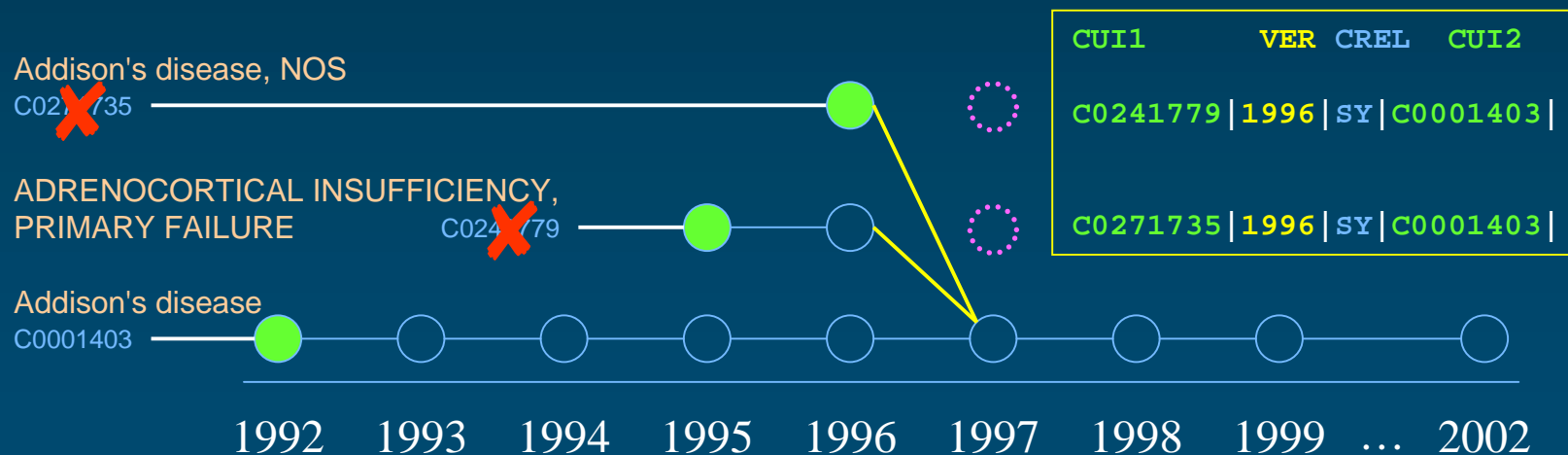
Metathesaurus Relationships

◆ Symbolic relations: ~5 M pairs of concepts

◆ Statistical relations :   ~6.5 M pairs of concepts
   (co-occurring concepts)

◆ Categorization: Relationships between concepts
   and semantic types from the Semantic Network

Symbolic relations

◆ Relation   MRREL

- Pair of concept identifiers
- Type
- Attribute (if any)
- List of sources (for type and attribute)

◆ Semantics of the relationship:
defined by its type [and attribute]

Symbolic relationships Type

- Hierarchical
  - Parent / Child   **PAR/CHD**
  - Broader / Narrower than   **RB/RN**
- Derived from hierarchies
  - Siblings (children of parents)   **SIB**
- Associative
  - Other   **RO**
- Various flavors of near-synonymy
  - Similar   **RL**
  - Source asserted synonymy   **SY**
  - Possible synonymy   **RQ**
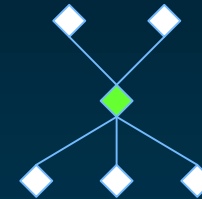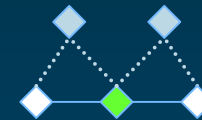
Symbolic relationships — Attribute

- ◆ Hierarchical
  - ● isa (is-a-kind-of)
  - ● part-of
- ◆ Associative
  - ● location-of
  - ● caused-by
  - ● treats
  - ● …
- ◆ Cross-references (mapping)

Addison's disease Hierarchical relations

Metathesaurus files Relations

◆ **Symbolic relations**   MRREL

◆ **Statistical relations**   MRCOC

◆ **Categorization**   MRSTY

> MRCXT is *not* the authoritative source of relationships

Semantic Types

Anatomical Structure

Fully Formed Anatomical Structure

Embryonic Structure

Disease or Syndrome

Body Part, Organ or Organ Component

Pharmacologic Substance

Population Group

Semantic Network

Metathesaurus

Medias-tinum

4

Saccular Viscus

Angina Pectoris

97

Esophagus

12

Heart

Cardiotonic Agents

225

Left Phrenic Nerve

Tissue Donors

22

Heart Valves

9

Fetal Heart

31
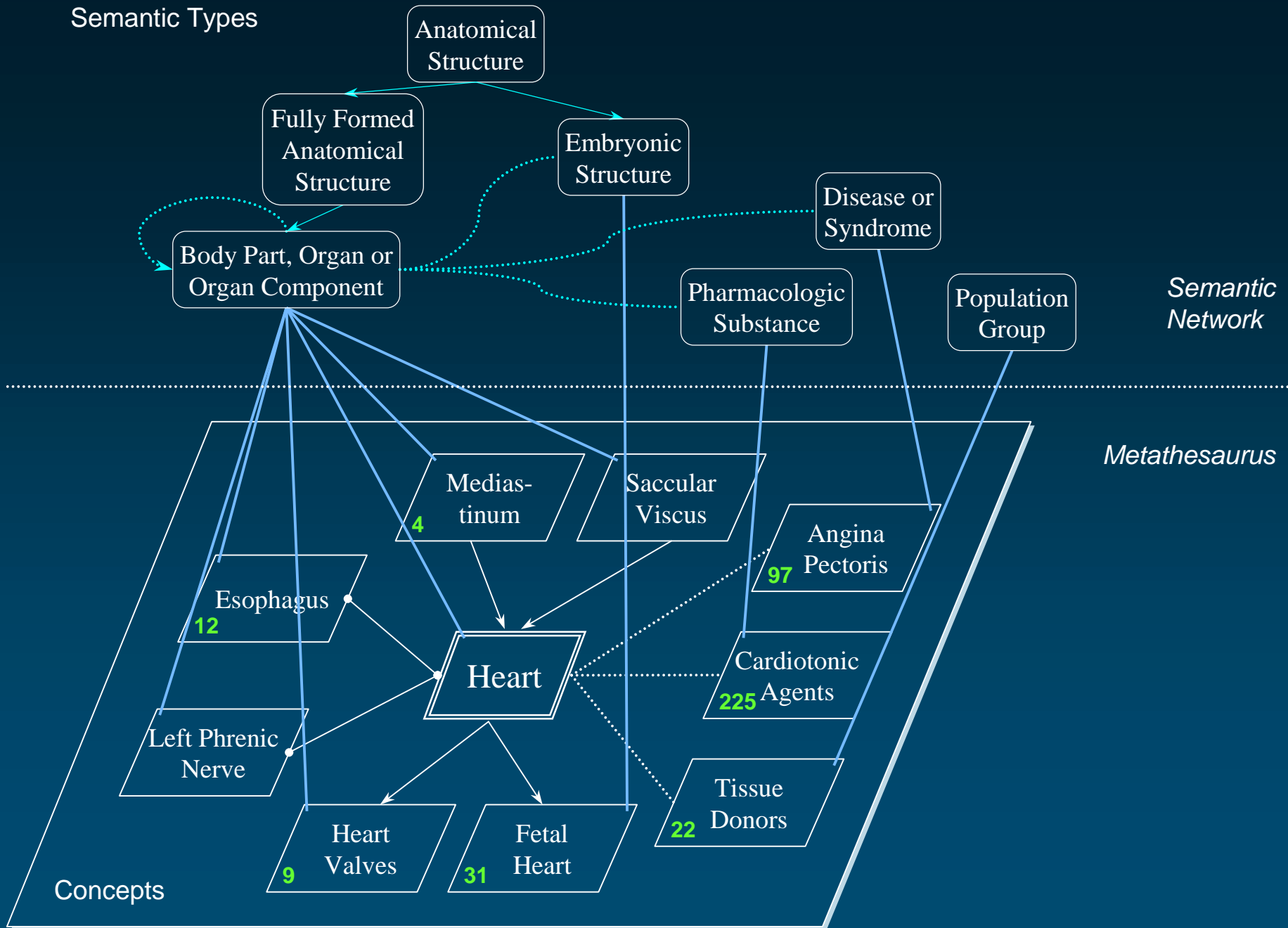
Concepts

UMLS Semantic Network

Semantic Network

- ◆ Semantic types (135)
  - ● tree structure
  - ● 2 major hierarchies
    - ■ Entity
      - – Physical Object
      - – Conceptual Entity
    - ■ Event
      - – Activity
      - – Phenomenon or Process

NLM

Semantic Network

- ◆ Semantic network relationships (54)
  - ● hierarchical (isa = is a kind of)
    - ▪ among types
      - – Animal *isa* Organism
      - – Enzyme *isa* Biologically Active Substance
    - ▪ among relations
      - – treats *isa* affects
  - ● non-hierarchical
    - ▪ Sign or Symptom *diagnoses* Pathologic Function
    - ▪ Pharmacologic Substance *treats* Pathologic Function

"Biologic Function" hierarchy (isa)
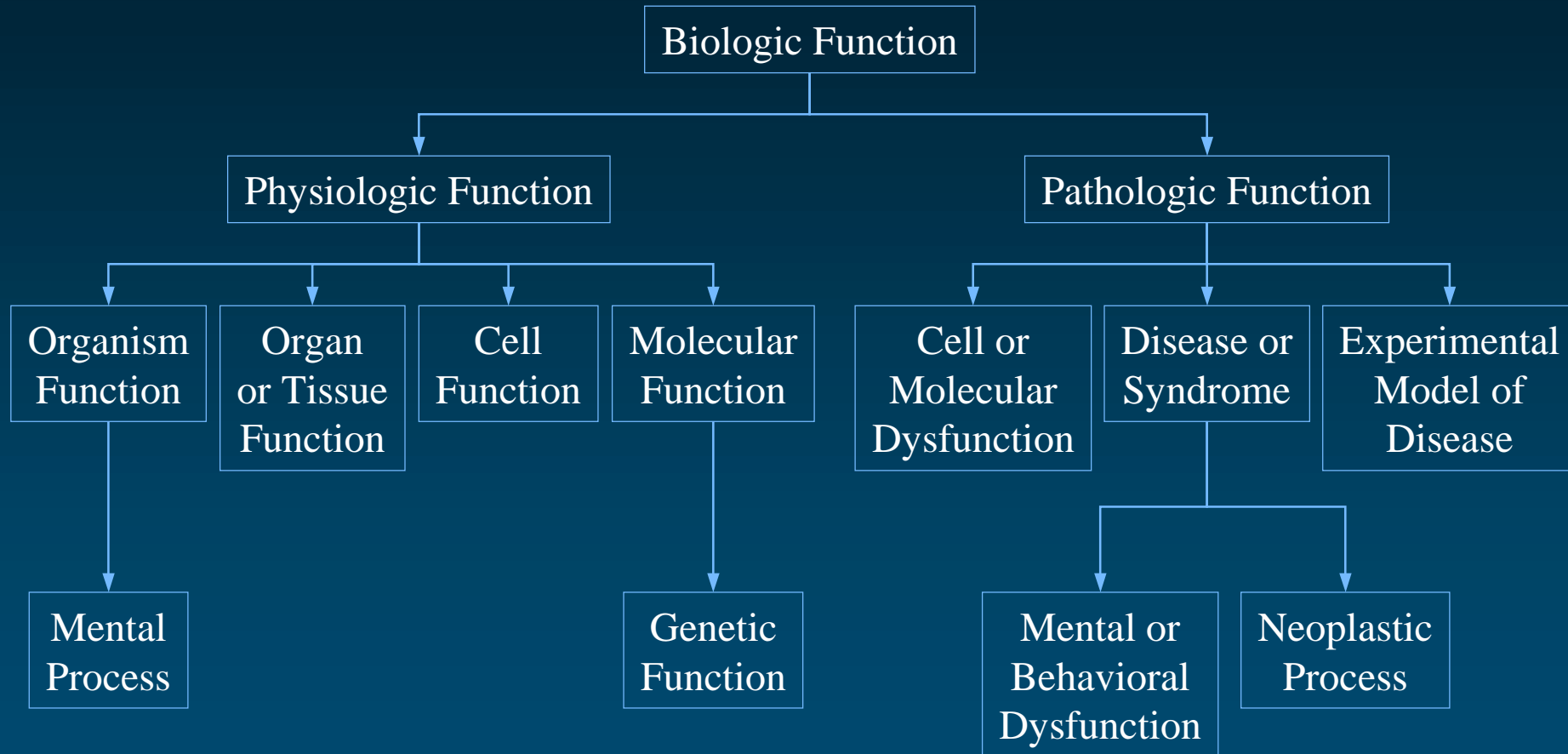
Associative (non-isa) relationships
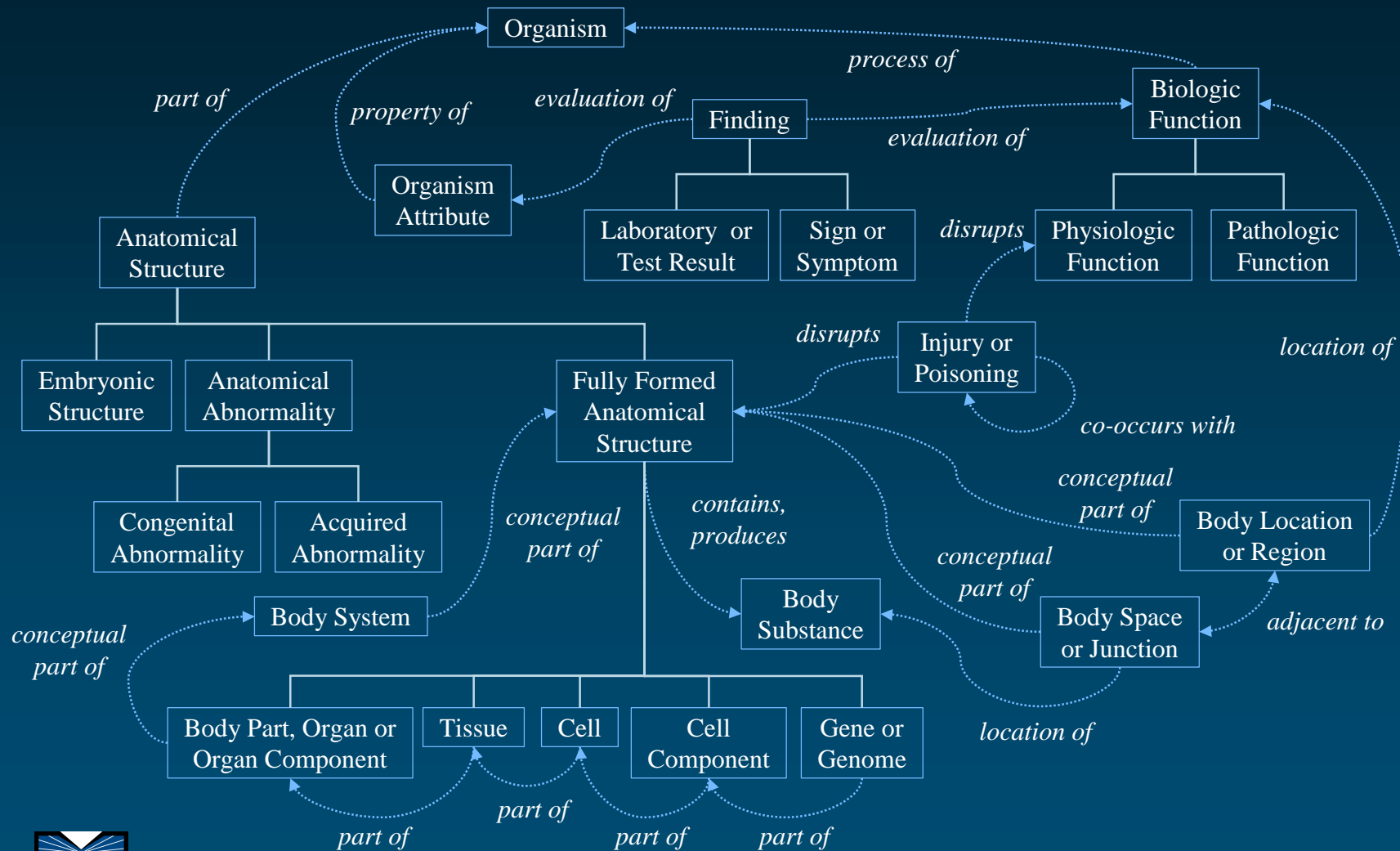
Role

- A relationship between 2 STs is a possible link between 2 concepts that have been assigned to those STs
  - The relationship may or may not hold at the concept level
  - Other relationships may apply at the concept level
- A child ST inherits properties from its parents (isa relationships)

Relationships can inherit semantics

Semantic Network

Fully Formed Anatomical Structure

location of

Biologic Function

isa

Pathologic Function

isa

Body Part, Organ, or Organ Component

isa

Disease or Syndrome

isa

Adrenal Cortex → Adrenal Cortical hypofunction

location of

Metathesaurus

65
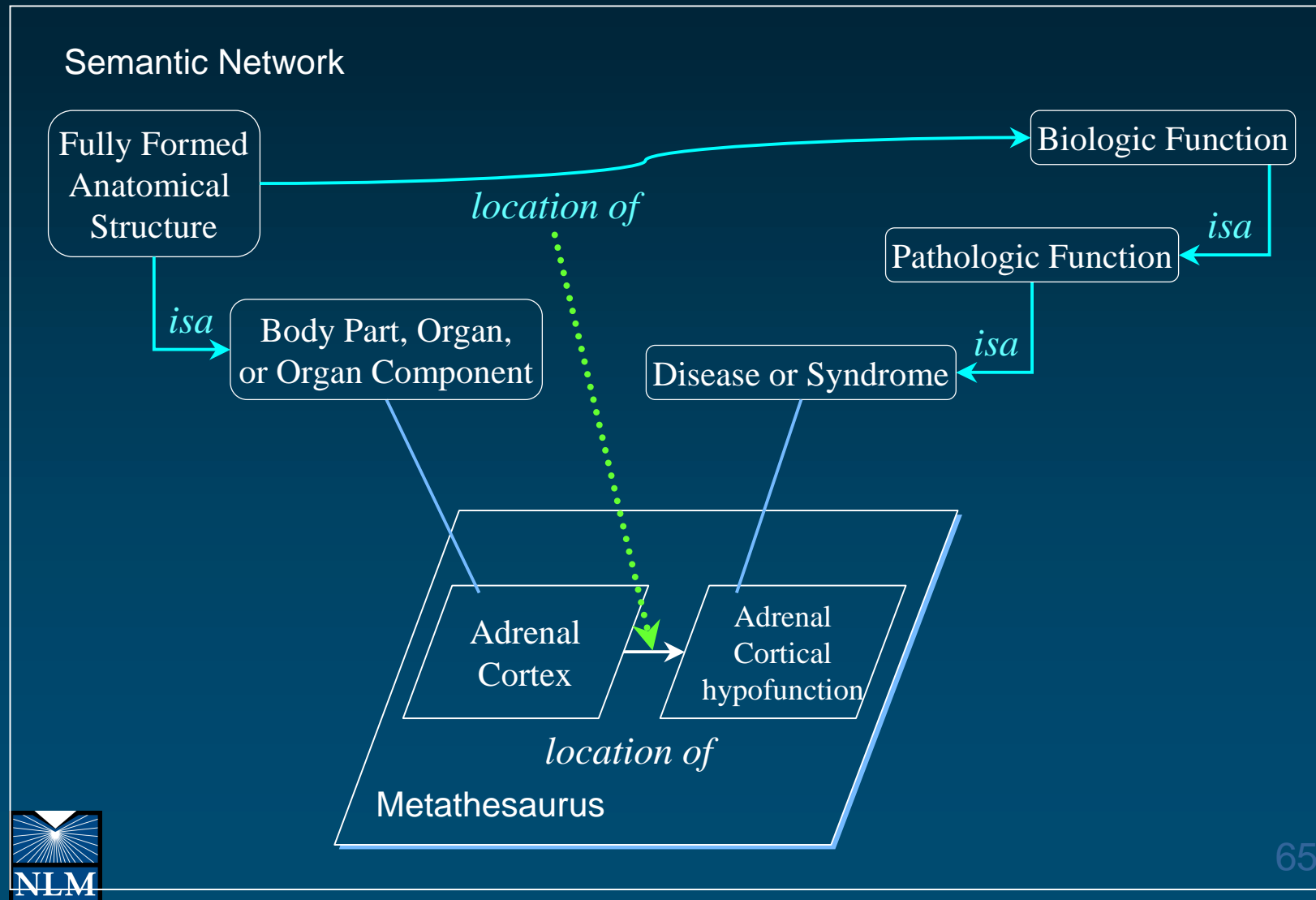
Applications

- ◆ To help qualify inter-concept relationships
  - using the relationships defined between their semantic types in the semantic network
- ◆ To strengthen the structure of the Metathesaurus
  - a relationship between 2 concepts should be consistent with the relationships defined between their semantic types in the semantic network
- ◆ Semantic interpretation
  - finding semantic relationships between concepts in text

SPECIALIST Lexicon
and lexical tools

SPECIALIST lexicon

- ◆ Content
  - English lexicon
  - Many words from the medical domain
- ◆ 160,000+ entries
- ◆ Word properties
  - morphology
  - orthography
  - syntax
- ◆ Used by the lexical tools

Morphology

- ◆ Inflection
  - ● noun   nucleus, nuclei
  - ● verb   cauterize, cauterizes, cauterized, cauterizing
  - ● adjective   red, redder, reddest
- ◆ Derivation
  - ● verb ⬌ noun   cauterize -- cauterization
  - ● adjective ⬌ noun   red -- redness

Orthography

- ◆ Spelling variants
  - ● oe/e   oesophagus - esophagus
  - ● ae/e   anaemia - anemia
  - ● ise/ize   cauterise - cauterize
  - ● genitive mark   Addison's disease
    Addison disease
    Addisons disease

Syntax

- ◆ Complementation
  - ● verbs
    - ▪ intransitive   I'll treat.
    - ▪ transitive   He treated the patient.
    - ▪ ditransitive   He treated the patient with a drug.
  - ● nouns
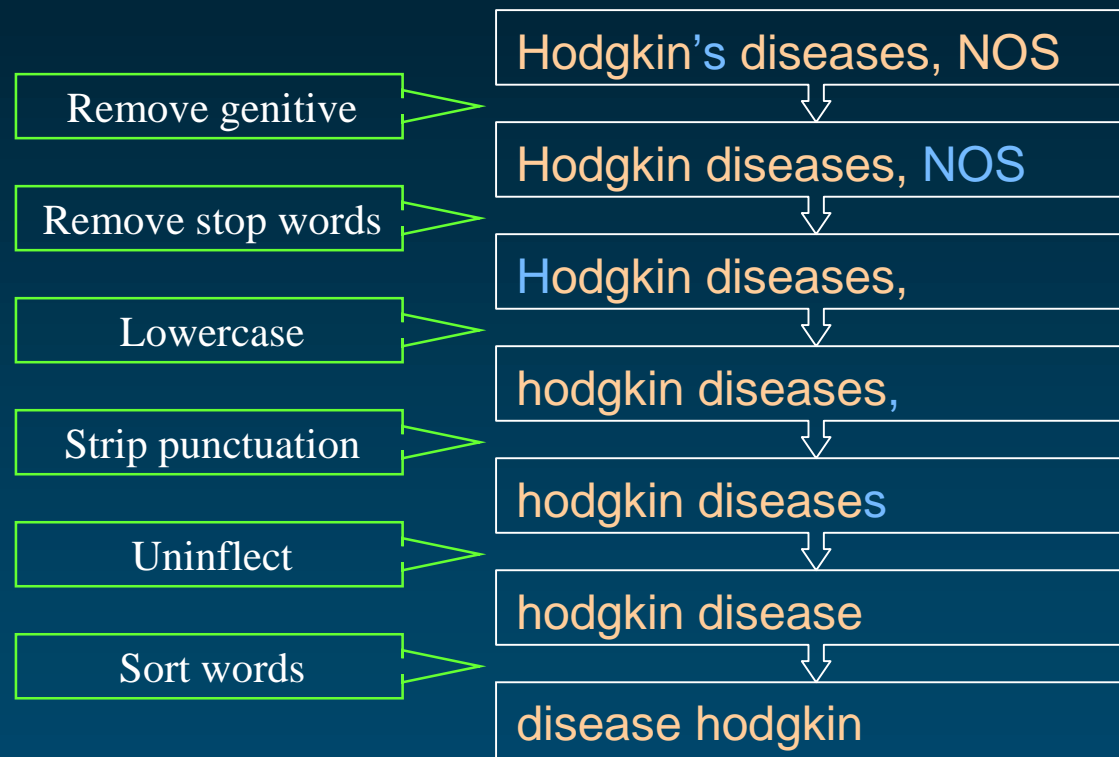   - ▪ prepositional phrase

        Valve of coronary sinus

- ◆ Position for adjectives

Lexical tools

- To manage lexical variation in biomedical terminologies
- Major tools
  - Normalization
  - Indexes
  - Lexical Variant Generation program (lvg)
- Based on the SPECIALIST Lexicon
- Used by noun phrase extractors, search engines

72

Normalization

Process		Text
		Hodgkin's diseases, NOS
Remove genitive	→	Hodgkin diseases, NOS
Remove stop words	→	Hodgkin diseases,
Lowercase	→	hodgkin diseases,
Strip punctuation	→	hodgkin diseases
Uninflect	→	hodgkin disease
Sort words	→	disease hodgkin

Normalization: Example

Hodgkin Disease
HODGKINS DISEASE
Hodgkin's Disease
Disease, Hodgkin's
Hodgkin's, disease
HODGKIN'S DISEASE
Hodgkin's disease
Hodgkins Disease
Hodgkin's disease NOS
Hodgkin's disease, NOS
Disease, Hodgkins
Diseases, Hodgkins
Hodgkins Diseases
Hodgkins disease
hodgkin's disease
Disease, Hodgkin

normalize ⟶ disease hodgkin

Normalization: Applications

- Model for lexical resemblance
- Help find lexical variants for a term
  - Terms that normalize the same usually share the same LUI
- Help find candidates to synonymy among terms
- Help map input terms to UMLS concepts

Indexes

- Word index
  - word to Metathesaurus strings
  - one word index per language
- Normalized word index
  - normalized word to Metathesaurus strings
  - English only
- Normalized string index
  - normalized term to Metathesaurus strings
  - English only

Lexical Variant Generation program

- Tool for specialists (linguists)
- Performs atomic lexical transformations
  - generating inflectional variants
  - lowercase
  - …
- Performs sequences of atomic transformations
  - a specialized sequence of transformations provides the normalized form of a term

Part II

How to use the UMLS?

Outline

◆ **Part II**

- Acquiring data and licensing mechanism
- Subsetting the Metathesaurus with *MetamorphoSys*
- Querying UMLS data
  - Relational tables and SQL queries
  - Object-oriented model and UMLS APIs
- UMLS-based applications *(MetaMap, Knowledge Source Server)*
- UMLS-based algorithms *(Restrict to MeSH)*
- Benefits and limitations

Acquiring data
and licensing mechanism

First step: License agreement

http://www.nlm.nih.gov/research/umls/license.html

◆ Sign and send to:

Sheldon Kotzin
Chief
Bibliographic Services Division
National Library of Medicine
8600 Rockville Pike
Bethesda, MD 20894
USA
Telephone 301-496-6217
Fax 301-496-0822
email kotzin@nlm.nih.gov

NOW THEREFORE, it is mutually agreed as follows:

1. The NLM hereby grants a nonexclusive, non-transferable right to LICENSEE to use the UMLS products and incorporate them in any computer applications or systems designed to improve access to biomedical information of any type subject to the restrictions in other provisions of this Agreement. The list of licensees authorized to use the UMLS products is public information.

2. No charges, usage fees or royalties will be paid to NLM. — *No charges to NLM*

3. LICENSEE is prohibited from distributing the UMLS products or subsets of these products, including individual vocabulary sources within the Metathesaurus®, except as an integral part of computer applications developed by LICENSEE for a purpose other than redistribution of data contained in the UMLS products. — *Do not redistribute*

4. LICENSEE agrees to inform NLM prior to distributing any application(s) in which it is using the UMLS products and is encouraged to inform NLM of any difficulties encountered in using the UMLS products, and changes or enhancements to the UMLS products that would make them more useful to LICENSEE and its user groups. — *Tell NLM how you use it (1)*

5. Within 30 days of the end of any calendar year in which LICENSEE makes use of the UMLS Metathesaurus, LICENSEE agrees to provide NLM with a brief report on the usefulness of the UMLS Metathesaurus in general and, if applicable, on the usefulness of CPT in the UMLS format in particular. — *Tell NLM how you use it (2)*

../..

NOW THEREFORE, it is mutually agreed as follows: (continued)

6. NLM represents that the data provided under this Agreement were formatted with a reasonable standard of care, but makes no warranties express or implied, including no warranty of merchantability or fitness for particular purpose, regarding the accuracy or completeness of the data or that the machine-readable copy is error free. Therefore, LICENSEE agrees to hold NLM, the Government, and any organizations contributing data to UMLS products free from any liability resulting from errors in data or on the machine-readable copy. NLM and all organizations contributing data to the UMLS products disclaim any liability for any consequences due to use, misuse, or interpretation of information contained or not contained in the UMLS products. *No warranty (1)*

7. NLM represents that its ability to continue to include certain vocabulary sources within the UMLS Metathesaurus is dependent on continuing contractual relations or agreements with the copyright holders for these vocabulary sources. Therefore, LICENSEE agrees to hold NLM free from any liability resulting from the removal of any vocabulary source from future editions of the UMLS Metathesaurus. *No warranty (2)*

8. NLM reserves the right to change the type and format of its machine-readable data. NLM agrees to inform LICENSEE of any changes to the format of UMLS data, EXCEPT the addition of entirely new data elements to the Metathesaurus, at least 90 days before the data are distributed. *The UMLS may change, but NLM will let you know*

9. The presence in UMLS products of data produced by organizations other than NLM does not imply any endorsement of the UMLS products by these organizations. *No endorsement by NLM*

../..

NOW THEREFORE, it is mutually agreed as follows: (continued)

10. Some of the Material in the UMLS Metathesaurus is from copyrighted sources. If LICENSEE uses any data from the UMLS Metathesaurus:

a) the LICENSEE is required to display in full, prior to providing user access to any Metathesaurus data, the following wording in order that its users be made aware of these copyright constraints:

Include this message

"Some material in the UMLS Metathesaurus is from copyrighted sources of the respective copyright claimants. Users of the UMLS Metathesaurus are solely responsible for compliance with any copyright restrictions and are referred to the copyright notices appearing in the original sources, all of which are hereby incorporated by reference."

and to display a list of all of the vocabularies obtained from the UMLS Metathesaurus that are used in the LICENSEE's application.

b) the LICENSEE is prohibited from altering data obtained from the UMLS Metathesaurus, but may include data from other sources in applications that also contain UMLS data. The LICENSEE may not imply in any way that data from other sources is part of the UMLS Metathesaurus or of any of its vocabulary sources.

Do not alter UMLS data

c) the LICENSEE is required to include in its applications identifiers from the UMLS Metathesaurus such that the original source vocabularies for any data obtained from the UMLS Metathesaurus can be determined by reference to a complete version of the UMLS Metathesaurus.

Include UMLS identifiers

../..

NOW THEREFORE, it is mutually agreed as follows: (continued)

11. LICENSEE shall acknowledge NLM as its source of the UMLS data, citing the year of the UMLS data, in a suitable and customary manner but may not in any way indicate or imply that NLM or any of the organizations whose vocabulary data are included in the UMLS has endorsed LICENSEE or its products.

Acknowledge NLM + specify version

12. For material in the UMLS Metathesaurus obtained from some sources additional restrictions on LICENSEE's use may apply. The categories of additional restrictions are described below. The list of UMLS Metathesaurus Vocabulary Sources, which is part of this Agreement and is updated annually, indicates the category of additional restrictions, if any, that apply to each vocabulary source.

4 categories of sources

LICENSEE should contact the copyright holder directly to discuss uses of a source vocabulary beyond those allowed under this license agreement. If LICENSEE or LICENSEE's end user has a separate agreement with the copyright holder for use of a UMLS Metathesaurus source vocabulary, LICENSEE or LICENSEE's end user may use data from that source obtained from the UMLS Metathesaurus in accordance with the terms of the separate agreement.

13. LICENSEE shall ensure that anyone who has authorized access to data from the UMLS Knowledge Sources under this Agreement complies with its provisions.

Make sure UMLS is protected in your applications

../..

NOW THEREFORE, it is mutually agreed as follows: (continued)

14. LICENSEE and/or its end users shall be solely responsible for compliance with any copyright or other restrictions on material in the UMLS Metathesaurus; NLM assumes no responsibility or liability associated with the LICENSEE's (or any of the LICENSEE's users) use and/or reproduction of copyrighted material. Anyone contemplating reproduction of all or any portion of any of the UMLS Metathesaurus should consult legal counsel.

15. This Agreement shall be effective until terminated by one of the parties upon 30 days written notice to the other party. LICENSEE's failure to abide by the terms of the Agreement shall be grounds for its termination. Neither the Government nor its employees shall be liable or responsible to LICENSEE in any manner whatsoever for damages of any nature whatsoever arising from the termination of this Agreement.

16. In the event that any provision of this Agreement is determined to violate any law or is unenforceable, the remainder of the Agreement shall remain in full force and effect.

License restriction levels

◆ Level 0 – 61.5% of concepts

 ● Basic license requirements, e.g., copyright statement and credits to NLM and producers of the vocabularies you use, no redistribution except as a part of your application

◆ Level 1 – 4.3% of concepts

 ● Basic, plus you must negotiate with producer to translate into another language

READ the license, including the appendix

License restriction levels

◆ Level 2 - .0009% of concepts

- Basic, plus you must negotiate with producer for use in the creation of health data

◆ Level 3 – 33.9% of concepts

- Basic, plus you must negotiate with the producer for *any* production use. Explicit prohibition against providing access via the Internet.

◆ There may - or may not - be license fees associated with uses not covered by the UMLS license.

Subsetting the Metathesaurus with *MetamorphoSys*

MetamorphoSys

- A tool distributed for use with the UMLS Knowledge Sources
  - Already present in UMLS distribution in $UMLSHOME/METAMSYS directory
- Multi-platform Java software
- Creates a customized version of the Metathesaurus

How does MetamorphoSys work?

| MRCON | MRSO | MRSTY | MRSAT | MRREL | [...] |

MetamorphoSys filter

| MRCON | MRSO | MRSTY | MRSAT | MRREL | [...] |

MetamorphoSys filter

| MRCON | MRSO | MRSTY | MRSAT | MRREL | [...] |
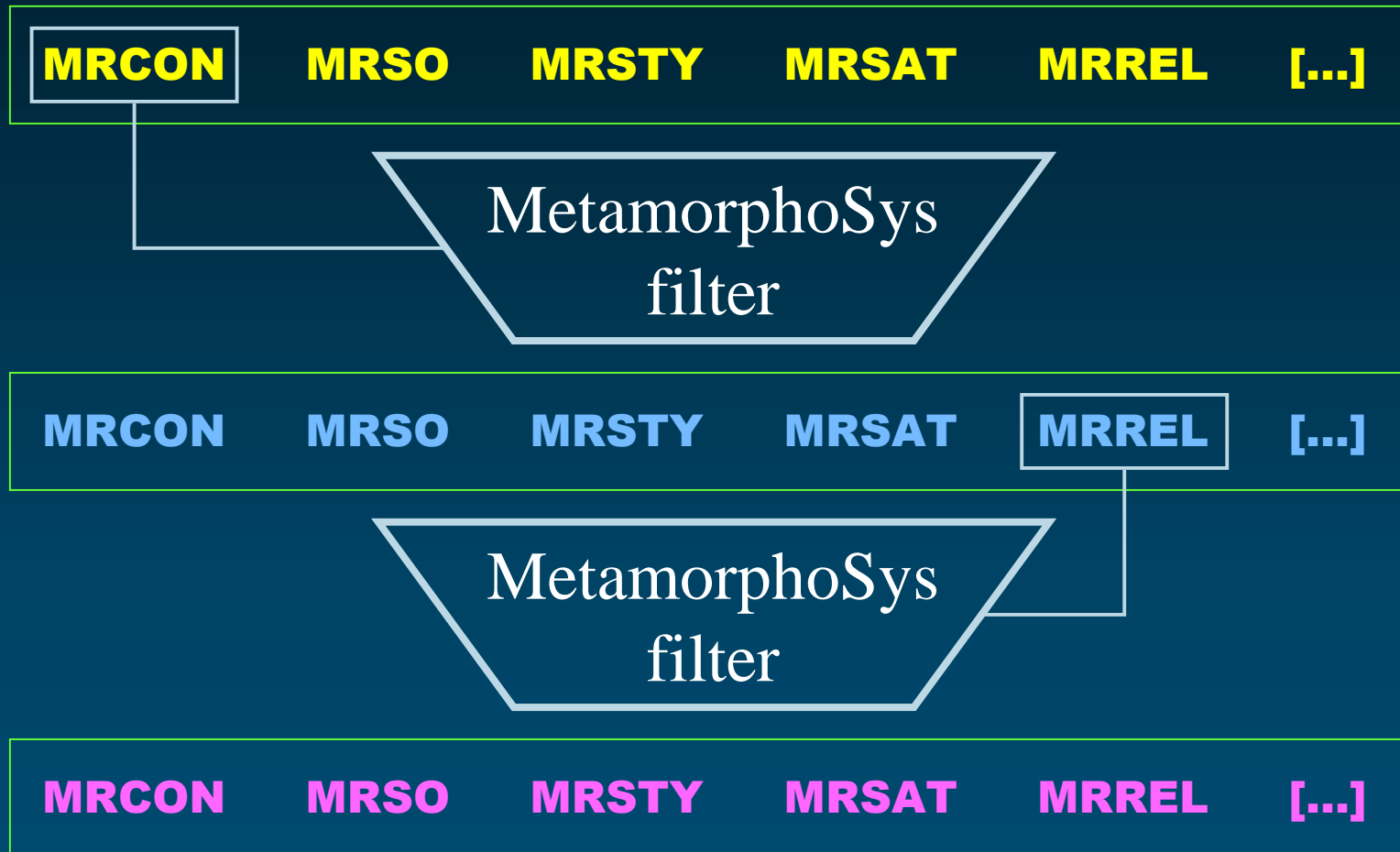
Filter by language

MRCON

Concept
C0001621

Term
L0001621
- **S0011232** *Adrenal Gland Diseases*
- **S0011231** Adrenal Gland Disease
- **S0000441** Disease of adrenal gland
- **S0481705** Disease of adrenal gland, NOS
- **S0220090** Disease, adrenal gland
- **S0044801** Gland Disease, Adrenal

[…]

Term
L0041793
- **S0860744** *Disorder of adrenal gland, unspecified*
- **S0217833** Unspecified disorder of adrenal glands

Term
L0161347
- **S0225481** *ADRENAL DISORDER*
- **S0627685** DISORDER ADRENAL (NOS)

[…]

Term
L0181041
- **S0632950** *Disorder of adrenal gland*
- **S0354509** Adrenal Gland Disorders

[…]

Term
L0368399
- **S0586222** *Adrenal disease*
- **S0466921** ADRENAL DISEASE, NOS

[…]

Term
Nebennierenkrankheiten   GER

Term
GLANDE SURRENALE, MALADIES   FRE   […]

92

Filter by source

MRSO

Concept
C0001621

Term L0001621	S0011232 *Adrenal Gland Diseases*	MeSH	
	S0011231 Adrenal Gland Disease	MeSH	
	S0000441 Disease of adrenal gland	SNOMED 2	
	~~S0000705 Disease of adrenal gland, NOS~~	~~SNOMED Intl~~	
	S0220090 Disease, adrenal gland	MeSH	
	S0044801 Gland Disease, Adrenal	MeSH	[...]

| Term
L0041793 | S0860744 *Disorder of adrenal gland, unspecified* | ICD-10 | |
| | S0217833 Unspecified disorder of adrenal glands | ICD-9 MedDRA | [...] |

| Term
L0161347 | S0225481 *ADRENAL DISORDER* | COSTAR CCPSS | |
| | S0627685 DISORDER ADRENAL (NOS) | COSTAR | [...] |

| Term
L0181041 | S0632950 *Disorder of adrenal gland* | CTV3 | |
| | S0354509 Adrenal Gland Disorders | Th. Psych | [...] |

| Term
L0368399 | S0586222 *Adrenal disease* | CTV3 | |
| | S0466921 ADRENAL DISEASE, NOS | COSTAR | [...] |

| Term
L1279026 | S1520972 *Nebennierenkrankheiten* | German MeSH | [...] |

| Term
L0162317 | S0226798 *SURRENALE, MALADIES* | French MeSH | [...] |

[...]

NLM

Filter by source

MRSO

Concept C0001621

Term L0001621		
S0011232 *Adrenal Gland Diseases*	MeSH	
S0011231 Adrenal Gland Disease	MeSH	
S0000441 Disease of adrenal gland	SNOMED 2	
S0481705 Disease of adrenal gland, NOS	SMOMED Intl	
S0220090 Disease, adrenal gland	MeSH	
S0044801 Gland Disease, Adrenal	MeSH	[...]

Term L0041793		
S0860744 *Disorder of adrenal gland, unspecified*	ICD-10	
S0217833 Unspecified disorder of adrenal glands	ICD-9 MedDRA	[...]

Term L0161347		
S0225481 *ADRENAL DISORDER*	COSTAR CCPSS	
S0627685 DISORDER ADRENAL (NOS)	COSTAR	[...]

Term L0181041		
S0000000 *Disorder of adrenal gland*	CTV3	
S0354509 Adrenal Gland Disorders	Th. Psych	[...]

Term L0368399		
S0000000 *Adrenal disease*	CTV3	
S0466921 ADRENAL DISEASE, NOS	COSTAR	[...]

Term L1279026		
S1520972 *Nebennierenkrankheiten*	German MeSH	[...]

Term L0162317		
S0226798 *SURRENALE, MALADIES*	French MeSH	[...]

[...]

NLM

94
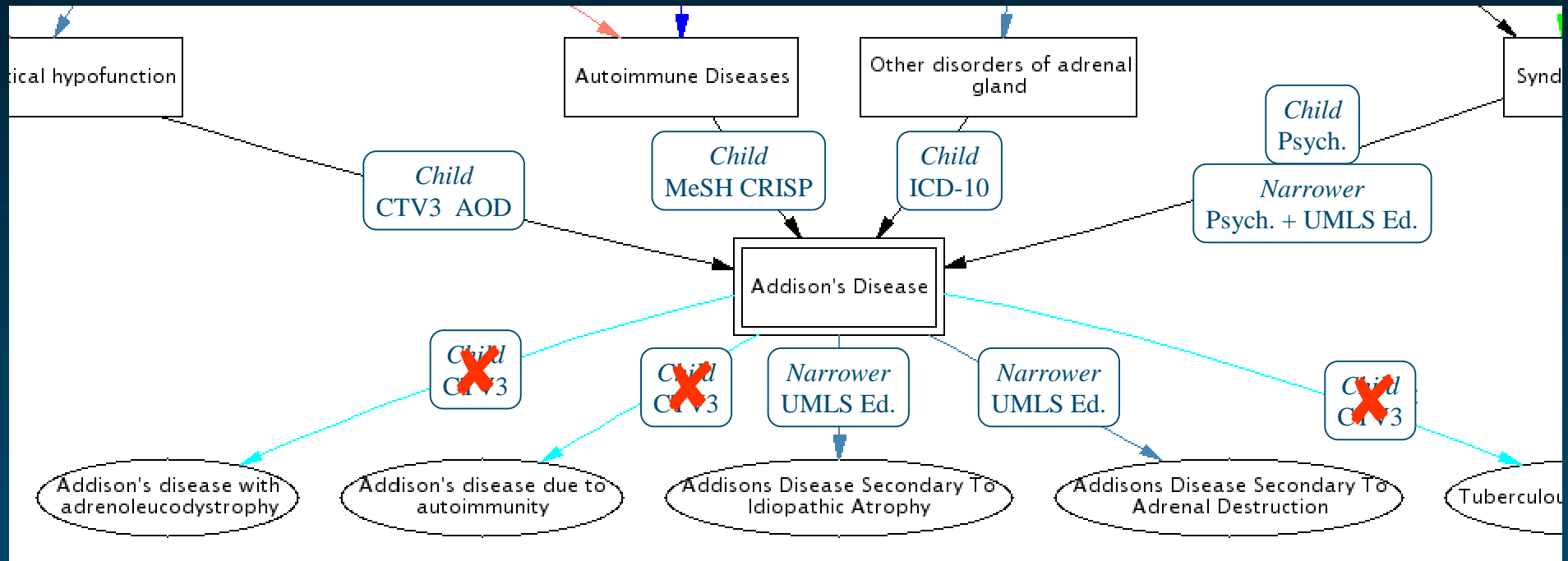
Exclude relationships

Exclude
Child in CTV3

MRREL

Other MetamorphoSys actions

◆ Modify precedence   **MRRANK**

◆ Exclude attribute   **MRSAT**

◆ Exclude suppressible strings

◆ Write your own filter

File   Edit   Options   Reset   Help

File/Folders | Sources To Exclude | Precedence | Term Status

Files and Folders

Installation Folder - Location of Metathesaurus Files

C:\small-META   Browse..

Target Folder - Location of Subset Files

C:\METASUBSET   Browse..

Current Configuration File

config/mmsys.prop.default   Browse..

CUI Input Handler

Efficient MR Files Input Stream   Browse..

CUI Output Handler

MR Files Versionless Output Stream   Browse..

File   Edit   Options   Reset   Help

File/Folders | Sources To Exclude | Precedence | Term Status

Please select one or more sources to remove from the UMLS Metathesaurus. For more info. on which categories of sources you might want to exclude consult the documentation. To select additional rows, hold down the <Cntrl> key while you make your selection. To reset selections to the default select "Reset Sources To Exclude Defaults" under the "Reset" menu.

Sources to Exclude

Full Source Name	Source Abbreviation	Source Family	Language	Level
AI/RHEUM	AIR93	AIR	ENG	0
Alternative Billing Concepts	ALT2000	ALT	ENG	3
Alcohol and Other Drugs Thesaurus	AOD2000	AOD	ENG	0
Beth Israel Vocabulary	BI98	BI	ENG	2
Portuguese translation of the Medical Subject Headings	BRMP2002	MSH	POR	3
Spanish translation of the Medical Subject Headings	BRMS2002	MSH	SPA	3
Canonical Clinical Problem Statement System	CCPSS99	CCPSS	ENG	3
Clinical Classifications Software	CCS99	CCS	ENG	0
Current Dental Terminology (CDT)	CDT3	HCPCS	ENG	3
COSTAR 1989	COS89	COS89	ENG	0
COSTAR 1992	COS92	COS92	ENG	0
COSTAR 1993	COS93	COS93	ENG	0
COSTAR 1995	COS95	COS95	ENG	0
Medical Entities Dictionary	CPM93	CPM	ENG	2
Physicians' Current Procedural Terminology, Spanish Translati...	CPT01SP	CPT	SPA	3
Physicians' Current Procedural Terminology	CPT2002	CPT	ENG	3
CRISP Thesaurus	CSP2002	CSP	ENG	0
COSTART	CST95	CST	ENG	0

File   Edit   Options   Reset   Help

| File/Folders | Sources To Exclude | Precedence | Term Status |

Select one or more source and term type combinations that you wish to make suppressible. To select additional rows hold down the <Cntrl> key while you make your selection. To reset selections to the default select "Reset Term Status Table Defaults" under the "Reset" menu.

Select One or More Suppressible Term Types

Source	Source Abbreviation	Term Type
ICD-9-CM. 6th ed.	ICD2002	HT
ICD-9-CM. 6th ed.	ICD2002	PT
International Classification of Primary Care	ICPC93	CC
International Classification of Primary Care	ICPC93	CO
International Classification of Primary Care	ICPC93	CP
International Classification of Primary Care	ICPC93	CS
International Classification of Primary Care	ICPC93	CX
International Classification of Primary Care	ICPC93	HT
International Classification of Primary Care	ICPC93	PC
International Classification of Primary Care	ICPC93	PS
International Classification of Primary Care	ICPC93	PT
International Classification of Primary Care	ICPC93	PX
ICPC, Basque Translation	ICPCBAQ	CP
ICPC, Basque Translation	ICPCBAQ	PT
ICPC, Danish Translation	ICPCDAN	CP
ICPC, Danish Translation	ICPCDAN	PT
ICPC, Dutch Translation	ICPCDUT	CP
ICPC, Dutch Translation	ICPCDUT	PT

File Edit Options Reset Help

Languages To Exclude | **Relationship Types To Exclude** | **Semantic Types To Exclude**

File/Folders | **Sources To Exclude** | **Precedence** | **Term Status** | **Attributes To Exclude**

Please select one or more attribute types to remove from the UMLS Metathesaurus.

Attributes to Exclude

Source	Source Abbreviation	Attribute Name
Alcohol and Other Drugs Thesaurus	AOD2000	HN
Alcohol and Other Drugs Thesaurus	AOD2000	SOS
Clinical Classifications Software	CCS99	CCI
CRISP Thesaurus	CSP2002	DID
CRISP Thesaurus	CSP2002	EZ
HCFA Common Procedure Coding System	HCPCS02	HAB
HCFA Common Procedure Coding System	HCPCS02	HAC
HCFA Common Procedure Coding System	HCPCS02	HAD
HCFA Common Procedure Coding System	HCPCS02	HAQ
HCFA Common Procedure Coding System	HCPCS02	HBT
HCFA Common Procedure Coding System	HCPCS02	HCC
HCFA Common Procedure Coding System	HCPCS02	HCD
HCFA Common Procedure Coding System	HCPCS02	HIR
HCFA Common Procedure Coding System	HCPCS02	HLC
HCFA Common Procedure Coding System	HCPCS02	HMP
HCFA Common Procedure Coding System	HCPCS02	HMR

NLM

File   Edit   Options   Reset   Help

| Languages To Exclude | Relationship Types To Exclude | Semantic Types To Exclude |

| File/Folders | Sources To Exclude | Precedence | Term Status | Attributes To Exclude |

Please select one or more languages to remove from the UMLS Metathesaurus.

Languages to Exclude

Language	Language Abbreviation
Basque	BAQ
Danish	DAN
Dutch	DUT
English	ENG
Finnish	FIN
French	FRE
German	GER
Hebrew	HEB
Hungarian	HUN
Italian	ITA
Norwegian	NOR
Portuguese	POR
Russian	RUS
Spanish	SPA
Swedish	SWE

Progress Monitor

◆ Once subsetting begins, a progress monitor tracks process

- Tracks progress through three major steps
- Screen disappears only when subsetting is complete
- "Cancel" ends the subsetting process
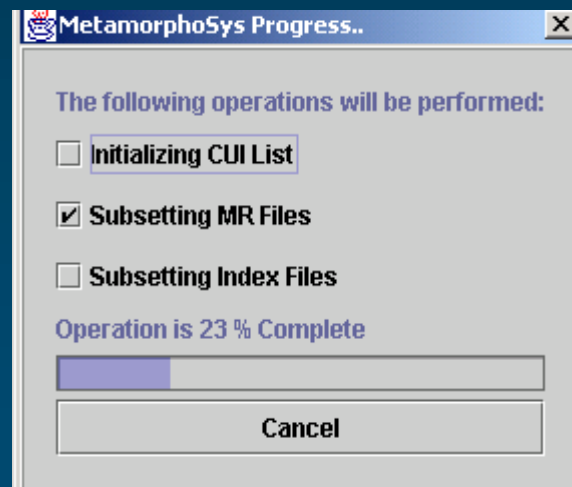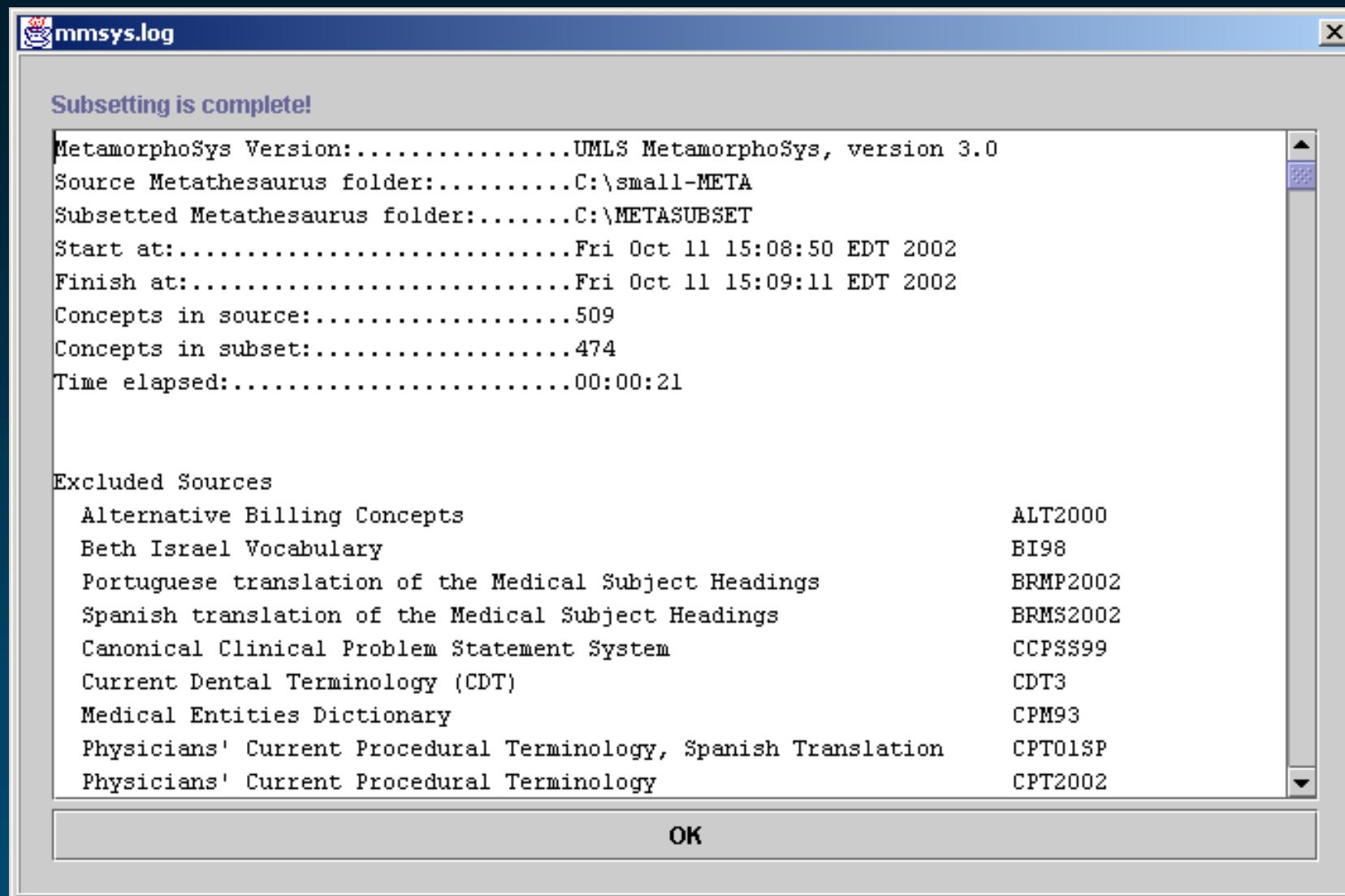
For More MetamorphoSys Information

◆ **UMLSinfo web site**   http://umlsinfo.nlm.nih.gov

  ● UMLS Tools section

◆ **UMLS Documentation**

  ● Section 2.8

Querying UMLS data (I):

Relational tables
and SQL queries

Creating a local UMLS database

◆ Load scripts   http://umlsinfo.nlm.nih.gov
  ● for MySQL, Oracle, and MS SQL server

```
/* Table: MRCON, records: 2097016, bytes: 147662196 */
DROP TABLE IF EXISTS MRCON\p\g
CREATE TABLE MRCON (
        CUI VARCHAR (8) BINARY NOT NULL,
        LAT VARCHAR (3) BINARY NOT NULL,
        TS VARCHAR (1) BINARY NOT NULL,
        LUI VARCHAR (8) BINARY NOT NULL,
        STT VARCHAR (3) BINARY NOT NULL,
        SUI VARCHAR (8) BINARY NOT NULL,
        STR BLOB NOT NULL,
        LRL VARCHAR (1) BINARY NOT NULL )\p\g
LOAD DATA LOCAL INFILE "../2002AD/META/MRCON" INTO TABLE MRCON FIELDS
TERMINATED BY "|"\p\g Select CURRENT_DATE, CURRENT_TIME\g
ALTER TABLE MRCON ADD INDEX MRCON_CUI_X (CUI), ADD INDEX MRCON_SUI_X
(SUI), ADD INDEX MRCON_STR_X (STR(6))\p\g
```

Simplified EA diagram

MRREL
- CUI1
- CUI2
- REL
- RELA
- SAB
- SRL

MRCOC
- CUI1
- CUI2
- SOC
- COT
- COF
- COA

MRSTY
- CUI
- TUI
- STY

MRCON
- CUI
- LUI
- SUI
- STR

MRSO
- CUI
- LUI
- SUI
- SAB
- TTY

MRDEF
- CUI
- SAB
- DEF

SRDEF
- RT
- UI
- STY/RL
- STN/RTN
- DEF

SRSTR
- STY/RL
- RL
- STY/RL
- LS

NLM

Sample query (1) Concepts by string

```
Select CUI, LUI, SUI, STR
From MRCON
Where STR like '%prostate%'
And LAT = 'ENG'
And TS <> 's'
And STT = 'PF'
```

- Avoid suppressible synonyms
- Consider using MetaMap instead

Sample query (2) Concept sources

```
Select MRCON.CUI, MRCON.CUI, MRCON.SUI,
STR, SAB, SCD
From MRCON, MRSO
Where MRCON.CUI = 'C0001403'
And MRCON.CUI = MRSO.CUI
And MRCON.LUI = MRSO.LUI
And MRCON.SUI = MRSO.SUI
```

- ◆ Join key = CUI + LUI + SUI

Sample query (3) Concepts by sem. type

```
Select CUI, LUI, SUI, STR
From MRCON, MRSTY
Where STY = 'Disease or Syndrome'
And MRCON.CUI = MRSTY.CUI
And LAT = 'ENG'
And STT = 'PF'
And TS = 'P'
```

◆ Join key = CUI only

◆ Consider using MetaMap instead

114

Querying UMLS data (II):

*Object-oriented model
and UMLS APIs*

http://umlsks4.nlm.n...sessionid=1pl8m0gys1

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Metathesaurus   Semantic Network   SPECIALIST Lexicon

Home

Documentation   Resources   Views/Profiles   Logout

About This Guide

This guide describes the installation, usage, and programmatic interface for the UMLS Knowledge Source Server (UMLSKS) Version 3.0.

Audience

The audience for this guide is the following set of people:

- UMLSKS System Administrators who will be installing the UMLSKS,
- Users of the UMLSKS web and command-line interfaces, and
- Developers of UMLSKS applications using the UMLSKS API.

Release Notes

The UMLSKS API Version 3.0 permits access to Metathesaurus data contained within the UMLS. Future versions of the API will allow access to the Semantic Network and Lexicon functions.

Version 3.0 of the UMLSKS provides a Java interface to the UMLSKS, as well as a TCP socket-based implementation of the UMLSKS for non-Java developers. A detailed list of features, bug fixes, and known problems can be found here.

How to Use This Guide

This manual contains the following chapters:

- Chapter 1 - Introduction describes the basic features and architecture of the UMLSKS.
- Chapter 2 - Installing the UMLSKS provides administrators instructions on installing and tailoring a UMLSKS installation.
- Chapter 3 - Using the Web Interface provides users with detailed instructions on interfacing with the UMLSKS using a standard web browser (Netscape, Internet Explorer, etc.).
- Chapter 4 - Using the Command-line Interface provides users with detailed instructions on interfacing with the UMLSKS using a web browser and a set of text commands.
- Chapter 5 - Building UMLSKS Software Applications describes the functions available to developers wanting to interface to the UMLSKS through another Java program.
- Chapter 6 - Using the XML Query Facility describes how to use the querying facility of the UMLSKS wherein users build XML queries to be executed.
- Chapter 7 - Using the UMLSKS Socket Server describes how to use the socket server to pass XML formatted commands or command-line type queries (e.g. ks -meta -c aids) that are to be executed by the server, with the results passed back to a client listening to the socket.

If You Need Help...

If you run into problems with the UMLSKS, please first consult the User's Guide for detailed instructions on usage and known problems with the system. Should you still encounter problems, please send a detailed message via email to the development team at the following email address: umlsks@nlm.nih.gov. Be sure to include the release version number for the UMLSKS system with your mailing.

User's Guide Contents

Print/Text Version

Table of Contents

About This Guide

Chapters

1. Introduction

2. Installing the UMLSKS

3. Using the Web Interface

4. Using the Command-line Interface

5. Building UMLSKS Software Applications

6. Using the XML Query Facility

7. Using the UMLSKS Socket Server

Document: Done (0.61 secs)

KSS API basics

◆ Remote server running at NLM

◆ Local application connected through

- Java RMI (Java-based applications)

    ▪ User guide: Chapter 5

    ▪ Java classes (part of the UMLS distribution)

- TCP/IP socket (XML-based queries)

    ▪ User guide: Chapter 7

    ▪ Socket server

        – Host: umlsks.nlm.nih.gov

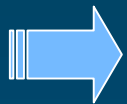   – Port: 8042

Sample query (1) Current version

```
<?xml version="1.0"?>
<getCurrentUMLSVersion version="1.0"/>
```
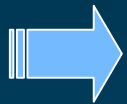
```
<?xml version="1.0"?>
<CurrentUMLSYear version="1.0">
    2003AA
</CurrentUMLSYear>
```

Sample query (2) Concepts by string

```xml
<?xml version="1.0"?>
<findCUI version="1.0">
<conceptName>prostate</conceptName>
<language>ENG</language>
<exact/>
<noSuppressibles/>
</findCUI>
```
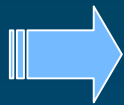
```xml
<?xml version="1.0"?>
<ConceptIdCollection version="1.0">
  <release>2003AA</release>
   <conceptId>
      <cui>C0033572</cui>
      <cn>Prostate</cn>
   </conceptId>
</ConceptIdCollection>
```

119

Sample query (3) Concepts properties

```xml
<?xml version="1.0"?>
<getSemanticType version="1.0">
<cui>C0033572</cui>
</getSemanticType>
```
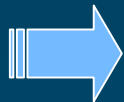
```xml
<?xml version="1.0"?>
<SemanticTypeCollection version="1.0">
<release>2003AA</release>
<cui>C0033572</cui>
<cn>Prostate</cn>
      <semanticType>
          <tui>T023</tui>
          <sty>Body Part, Organ,
              or Organ Component</sty>
      </semanticType>
</SemanticTypeCollection>
```
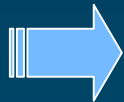
Sample query (4) Relationships

```
<?xml version="1.0"?>
<getRelations version="1.0">
<cui>C0033572</cui>
<rel>RO</rel>
</getRelations>
```

```
<?xml version="1.0"?>
<RelationCollection version="1.0">
[…]
        <relation>
            <rel>RO</rel>
            <cui2>C0005001</cui2>
            <cn2>Prostatic Hypertrophy, Benign</cn2>
            <rela>has_location</rela>
            <sab>SNMI</sab>
            <sl>SNMI</sl>
            <mg></mg>
        </relation>
[…]
```

121

Sample query (5) All semantic type Ids

```
<?xml version="1.0"?>
<listSemTypeIds version="1.0">
</listSemTypeIds>
```

```
<?xml version="1.0"?>
<SemNetIdCollection version="1.0">
     <release>2003AA</release>
     <semnetId>
         <name>Organism</name>
         <ui>T001</ui>
         <semtype/>
     </semnetId>
     <semnetId>
         <name>Plant</name>
         <ui>T002</ui>
         <semtype/>
     </semnetId>
     […]
```

122

UMLS-based applications

MetaMap
Knowledge Source Server

MetaMap

MetaMap Motivation

- ◆ **Information extraction**
  - ● Identifying UMLS concepts from text
- ◆ Usage
  - ● Information indexing and retrieval
  - ● Knowledge extraction / discovery
  - ● Semantic interpretation
- ◆ Characteristics
  - ● Linguistic approach
  - ● Based on UMLS knowledge sources

125

MetaMap Methods

- Parsing
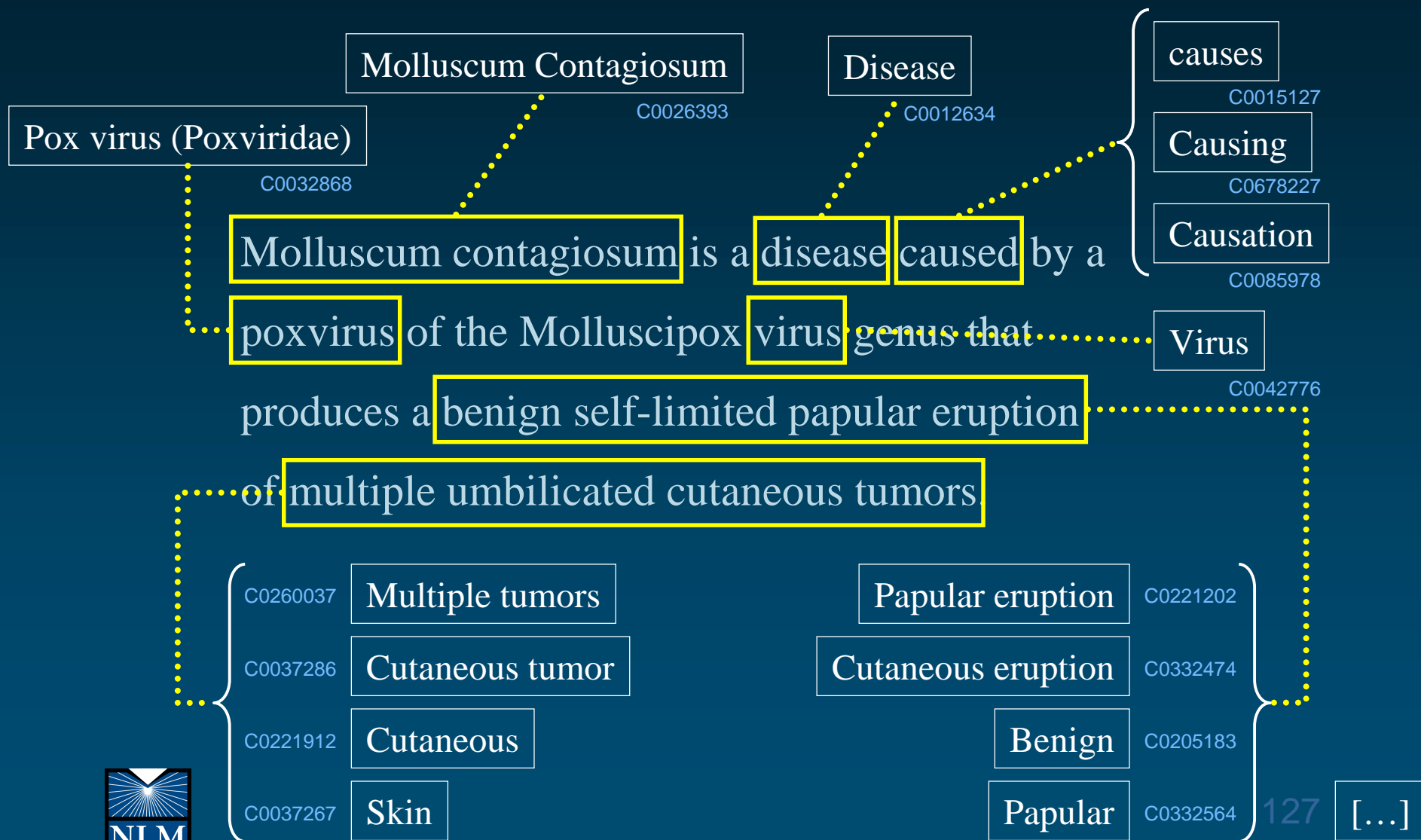 - Shallow syntactic analysis
  - SPECIALIST lexicon
  - Xerox part-of-speech tagger
- Variant generation
- Candidate retrieval
  - Retrieve candidate terms containing at least one variant
- Candidate evaluation
  - Rank candidate terms with respect to closeness to input text (centrality, variation, coverage, and cohesiveness)

MetaMap Example

Molluscum Contagiosum
C0026393

Disease
C0012634

causes
C0015127

Pox virus (Poxviridae)
C0032868

Causing
C0678227

Causation
C0085978

Molluscum contagiosum is a disease caused by a poxvirus of the Molluscipox virus genus that produces a benign self-limited papular eruption of multiple umbilicated cutaneous tumors

Virus
C0042776

C0260037 Multiple tumors

C0037286 Cutaneous tumor

C0221912 Cutaneous

C0037267 Skin

Papular eruption C0221202

Cutaneous eruption C0332474

Benign C0205183

Papular C0332564

127 […]

Semantic Network

Virus

location of

causes

Body Part, Organ,
or Organ Component

Pathologic
Function

Disease or
Syndrome

manifes-tation of

Neoplastic
Process

Finding

Virus

Disease

Pox virus
(Poxviridae)

Papular eruption

Molluscum
Contagiosum

Multiple
tumors

Cutaneous
tumor

Skin

Cutaneous
eruption

Metathesaurus

Using MetaMap MMTx

- Requires UMLS license
- Local implementation (Java-based)
- Provides
  - Stand-alone application
  - API for integrating in other applications

http://mmtx.nlm.nih.gov

Knowledge Source Server

UMLS Knowledge Source Server (UMLSKS)

Welcome to the NLM's Unified Medical Language System Knowledge Source Server (UMLSKS)

The UMLSKS is a set of web based interaction tools and a programmer interface to allow users and developers access to the biomedical terminologies found within the UMLS. This site is the starting point for access to the three knowledge sources developed and distributed by the National Library of Medicine as part of the UMLS project.

**What's in the UMLSKS?**

UMLS Metathesaurus®
Semantic Network
SPECIALIST Lexicon

Available UMLS Documentation

UMLS Release 2003AA
UMLS Release 2002AD
UMLS Release 2002AC
UMLS Release 2002AB
UMLS Release 2002AA

What's New in Version 3.0?

- Release 3.0 New Features & Bug Fixes
(Last Updated 1/23/03)

- 2001, 2002AA, 2002AB, 2002AC, 2002AD, and 2003AA Data Available for Download and Search!

- Version 1.0 will remain available; however, it will be unsupported (i.e. new features will not be added to it, nor will bug fixes be made).

Have a Question?

Perhaps it's answered in the FAQ

UMLSKS Registered Users:

Login ID: bodenreider

Password: xxxxxxxx

[Login]

Forgotten your password?

New UMLSKS Users:

The UMLSKS is restricted to registered users. If you are not a registered user and would like to use the resources of the UMLS, please read the UMLS license agreement, obtain a UMLS registration number, and fill in the form below to request a UMLSKS user account.

Login ID:
Password:
Confirm Password:
First Name:
Last Name:
Email:
UMLS License #:
UMLS Licensee First Name:
UMLS Licensee Last Name:

[Request Account]

Users are responsible for compliance with the UMLS copyright restrictions.
Users are also encouraged to read the NLM Copyright and Privacy Policy.

U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Department of Health & Human Services

Users are responsible for compliance with UMLS copyright restrictions
Comments/Suggestions? Email umlsks@nlm.nih.gov with your input.

umlsks.nlm.nih.gov

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Metathesaurus   **Semantic Network**   **SPECIALIST Lexicon**

Home   Documentation   Resources   Views/Profiles   Logout

Welcome to the NLM's Unified Medical Language System Knowledge Source Server (UMLSKS)

The UMLSKS is a set of web based interaction tools and a programmer interface to allow users and developers access to the biomedical terminologies found within the UMLS. This site is the starting point for access to the three knowledge sources developed and distributed by the National Library of Medicine (NLM) as part of the UMLS project.

Metathesaurus®

The UMLS Metathesaurus® contains information about biomedical concepts and terms from many controlled vocabularies and classifications used in patient records, administrative health data, bibliographic and full-text databases, and expert systems.

Search
Advanced Search

Semantic Network
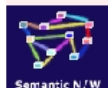
The Semantic Network, through its semantic types, provides a consistent categorization of all concepts represented in the UMLS Metathesaurus. The links between the semantic types provide the structure for the Network and represent important relationships in the biomedical domain.

Search
Advanced Search

SPECIALIST Lexicon

The SPECIALIST Lexicon is an English language lexicon with many biomedical terms, containing syntactic, morphological, and orthographic information for each term or word.

Search
Advanced Search

My Profile

The Views/Profiles segment of the UMLSKS allows users to modify their profile information, including their UMLSKS access information.

View/Edit

What's New?

New!

- Version 3.0 released on February 15, 2003! See the Release 3.0 New Features & Bug Fixes document for details on what's new.
- Having problems or have a question about the UMLSKS? Perhaps it's answered in the FAQ! Check back frequently for updates.
- Version 1.0 of the UMLSKS will remain available; however, it will be unsupported. Any new features will not be added to Version 1.0, nor will any bug fixes be made. Please update your software to the new version.

Available Documentation

Fact Sheet
Version 3.0 Release Notes
User's Guide
API Documentation
UMLS Documentation Set

UMLSKS Resources

Metathesaurus
Metathesaurus String Properties

Semantic Network
Semantic Navigator
Semantic Groups

NLP & Lexical Resources
Lexical Tools
MetaMap Transfer (MMTx)
Word Sense Disambiguation (WSD) Test Collection

Downloads

UMLS Knowledge Sources
UMLSKS API & Documentation

Sending request to umlsks4.nlm.nih.gov...

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Metathesaurus   **Semantic Network**   **SPECIALIST Lexicon**

Home   Search   Advanced Search   Documentation   Resources   Views/Profiles   Logout

Metathesaurus Search

Select UMLS Release: [2003AA ▶]

Enter a term or a concept unique identifier (CUI):   [addison's disease]

Perform Concept Search

Basic searching allows users to search for a concept that contains either a user entered term name or a concept unique identifier (CUI). The term name/CUI (e.g. C0001175) is entered into the text entry box above.

Pressing the *Perform Concept Search* button will update the display to show concepts found with the entered term name or with the requested CUI. Concept information is available during a concept search and includes synonyms, hierarchical context details, relations, and co-occurring concepts.

U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Department of Health & Human Services

Users are responsible for compliance with UMLS copyright restrictions
Comments/Suggestions? Email umlsks@nlm.nih.gov with your input.

http://umlsks4.nlm.n...sessionid=1pl8m0gys1

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Metathesaurus   **Semantic Network**   **SPECIALIST Lexicon**

Home   Search   Advanced Search   Documentation   Resources   Views/Profiles   Logout

Metathesaurus Search for: **addison's disease** in UMLS Release **2003AA**

Display   **Display All**

Concept
☑ Definition
☑ Synonyms
☐ Other Languages
☐ Suppressible Synonyms
☐ Sources

Context
☑ Ancestors
☐ Parents
☐ Siblings
☐ Children

Relations
☐ Narrower
☐ Broader
☐ Similar
☐ Other
☐ Related and possibly
synonymous
☐ Source asserted synonymy
☐ Allowable Subheadings
☐ Associated Expressions

Co-occurring Concepts
☐ Co-occurring MeSH
☐ Co-occurring MeSH By
Semantic Group
☐ Co-occurring AI/RHEUM
☐ Co-occurring AI/RHEUM By
Semantic Group

Concept: Addison's Disease

CUI: C0001403

Semantic Type: Disease or Syndrome

Definition:
A disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. In the absence of replacement therapy, it is usually fatal. (MeSH)

Synonyms:
Addison's Disease
Addison melanoderma
ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)
ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE
Asthenia pigmentosa
Bronzed disease
Melasma addisonii
Primary adrenal deficiency
Primary adrenal insufficiency
Primary adrenocortical insufficiency

Ancestors:

MeSH
  MeSH Descriptors []
    Index Medicus Descriptor []
      Diseases (MeSH Category) [C]
        Endocrine Diseases [C19]
          Adrenal Gland Diseases [C19.053]
            Adrenal Gland Hypofunction [C19.053.264]
              Addison's Disease [C19.053.264.263]

MeSH
  MeSH Descriptors []
    Index Medicus Descriptor []
      Diseases (MeSH Category) [C]
        Immunologic Diseases [C20]
          Autoimmune Diseases [C20.111]
            Addison's Disease [C20.111.163]

Alcohol and Other Drug Thesaurus
  health and disease [G]
    disorder by body system or organ function [GK]
      endocrine disorder [GV]
        adrenal disorder [GV14]

Document: Done (1.015 secs)

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Metathesaurus   **Semantic Network**   **SPECIALIST Lexicon**

Home   Search   Advanced Search Documentation   Resources   Views/Profiles   Logout

Metathesaurus Search for: **addison's disease** in UMLS Release **2003AA**

Display   **Display All**

Concept
- ☐ Definition
- ☐ Synonyms
- ☐ Other Languages
- ☐ Suppressible Synonyms
- ☐ Sources

Context
- ☐ Ancestors
- ☐ Parents
- ☐ Siblings
- ☐ Children

Relations
- ☐ Narrower
- ☐ Broader
- ☐ Similar
- ☐ Other
- ☐ Related and possibly synonymous
- ☐ Source asserted synonymy
- ☐ Allowable Subheadings
- ☐ Associated Expressions

Co-occurring Concepts
- ☑ Co-occurring MeSH
- ☐ Co-occurring MeSH By Semantic Group
- ☐ Co-occurring AI/RHEUM
- ☐ Co-occurring AI/RHEUM By Semantic Group

Concept: Addison's Disease

CUI: C0001403

Semantic Type: Disease or Syndrome

Co-occurring MeSH Terms:

[22] Autoantibodies
[21] Hydrocortisone
[19] Autoimmune Diseases
[15] Steroid 21-Hydroxylase
[14] Adrenal Glands
[14] Polyendocrinopathies, Autoimmune
[13] Glucocorticoids
[12] Adrenal Cortex
[11] Corticotropin
[11] Adrenoleukodystrophy
[10] Adrenal Gland Diseases
[10] Adrenal Gland Neoplasms
[10] Diabetes Mellitus, Insulin-Dependent
[8] Chronic Fatigue Syndrome
[8] Hyponatremia
[8] Tuberculosis
[8] Tuberculosis, Endocrine
[8] Hormone Replacement Therapy
[8] Chronic lymphocytic thyroiditis
[7] Bone Density
[7] Dog Diseases
[6] Adrenal gland hypofunction
[6] Anti-Inflammatory Agents, Steroidal
[6] Cushing Syndrome
[6] Prasterone
[6] Fludrocortisone
[6] Hypoglycemia
[5] Fatigue <1>
[5] Hyperkalemia
[5] Hypothyroidism
[4] Transaminases <1>
[4] Circadian Rhythm
[4] Esophageal Achalasia
[4] Hypercalcemia
[4] Polymorphism (Genetics)
[4] Pregnancy Complications
[4] Lymphoma, B-Cell
[4] Antiphospholipid Syndrome
[3] Adjuvants, Immunologic

Document: Done (1.75 secs)

Netscape

http://umlsks4.nlm.n...sessionid=1pl8m0gys1

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Metathesaurus   **Semantic Network**   **SPECIALIST Lexicon**

Home   Search   Advanced Search   Documentation   Resources   Views/Profiles   Logout

This version of the Semantic Network browser is based on a Java SWING applet. This is best viewed in Netscape 6.x and above or IE 5.x and above. If you have an older version of a browser please click _here_ for the HTML version of the Semantic Network browser.

Semantic Network

Semantic Types: Entity, Event   **Semantic Relations:** isa, associated with

Enter a Semantic Type or Relation in the box below or select from the list below and click on the Find button:

Semantic Types:
| Disease or Syndrome |

Semantic Relations:
| adjacent_to |

Find

Event
- Activity
- Phenomenon or Process
  - Injury or Poisoning
  - Human-caused Phenomenon or Process
  - Natural Phenomenon or Process
- Biologic Function
  - Physiologic Function
  - Pathologic Function
    - Disease or Syndrome
      - Cell or Molecular Dysfunction
      - Experimental Model of Disease

☐ Concepts   ☑ Relations   ☐ Raw records

Semantic Type: Disease or Syndrome

TUI: T047

Definition: A condition which alters or interferes with a normal process, state, or activity of an organism. It is usually characterized by the abnormal functioning of one or more of the host's systems, parts, or organs. Included here is a complex of symptoms descriptive of a disorder.

LHS Relations

Disease or Syndrome conceptually related to Experimental Model of Disease
Disease or Syndrome isa Pathologic Function
Disease or Syndrome produces Tissue

Applet gov.nlm.nlm.kss.gui.applet.SemanticNetwork started

http://umlsks4.nlm.n...sessionid=1pl8m0gvs1

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 3.0   UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA

Home

Metathesaurus   Semantic Network   SPECIALIST Lexicon

Documentation   Resources   Views/Profiles   Logout

The following documents are provided as resources for users of the UMLSKS.

UMLSKS User's Guide
---Last Updated: 01/15/03---

Provides instructions on installing the UMLSKS software and detailed instructions on using the UMLSKS Interface and Application Programmer Interface (API).

UMLSKS API Download and Documentation
---Last Updated: 01/15/03---

Provides developers with all of the software components necessary to build and run a client program that communicates with the UMLSKS.

2003, 2002 UMLS Documentation Set

Accesses the latest documentation set for Release Year 2002AA, 2002AB, 2002AC, 2002AD and 2003AA for the UMLS data, including the Metathesaurus, Semantic Network, and SPECIALIST Lexicon details.

- **UMLS Release 2003AA** *
  ---Last Updated: 12/23/02---

- **UMLS Release 2002AD** *
  ---Last Updated: 11/08/02---

- **UMLS Release 2002AC** *
  ---Last Updated: 10/23/02---

- **UMLS Release 2002AB** *
  ---Last Updated: 5/1/02---

- **UMLS Release 2002AA** *
  ---Last Updated: 1/29/02---

* UMLS CD-ROMs are no longer automatically mailed to licensed users, as the information is available here or via download. Those who wish the 2002 UMLS Knowledge Sources on CD-ROM should email their request to the UMLS Support Group. Be sure to include the UMLS license number and correct return mailing address with the request.

U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Department of Health & Human Services

Users are responsible for compliance with UMLS copyright restrictions
Comments/Suggestions? Email umlsks@nlm.nih.gov with your input.

Document: Done (0.469 secs)

UMLS-based algorithm

Restrict to MeSH

Indexing Initiative

- For noun phrases extracted from medical texts, map to UMLS concepts

- Then, select from the MeSH vocabulary the concepts that are the most closely related to the original concepts

Medical text

Noun phrase

MeSH descriptor

UMLS

140
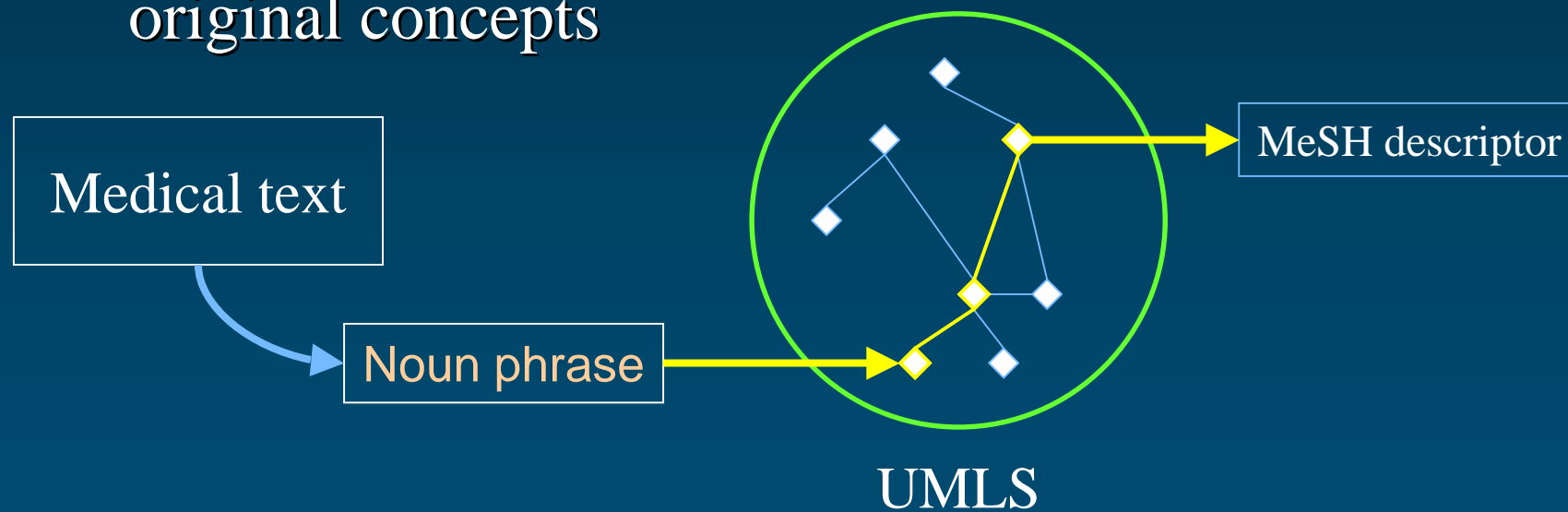
Restrict to MeSH [Bodenreider & al., *AMIA*, 1998]

◆ Based on the principle of semantic locality

◆ Use different components of the UMLS

◆ 4 techniques of increasing aggressiveness
  - Use Synonymy   MRCON + MRSO
  - Use Associated expressions (ATXs)   MRATX
  - Explore the Ancestors   MRREL + SN
 - Explore the Other related concepts   MRREL + SN

Restrict to MeSH Synonymy

- Term mapped to Source concept
- For this concept, is there a synonym term that comes from MeSH? (MRSO)

Restrict to MeSH Assoc. expressions

- ◆ If not,

- ◆ Is there an associated expression (ATX) that describes this concept using a combination of MeSH descriptors? (MRATX)

Endoscopic removal of intraluminal foreign body from oesophagus without incision ⇔

```
                    ( AND )
          ( MH/SH )
   [ Esophagus ]  [ surgery ]     [ Foreign Bodies ]
```
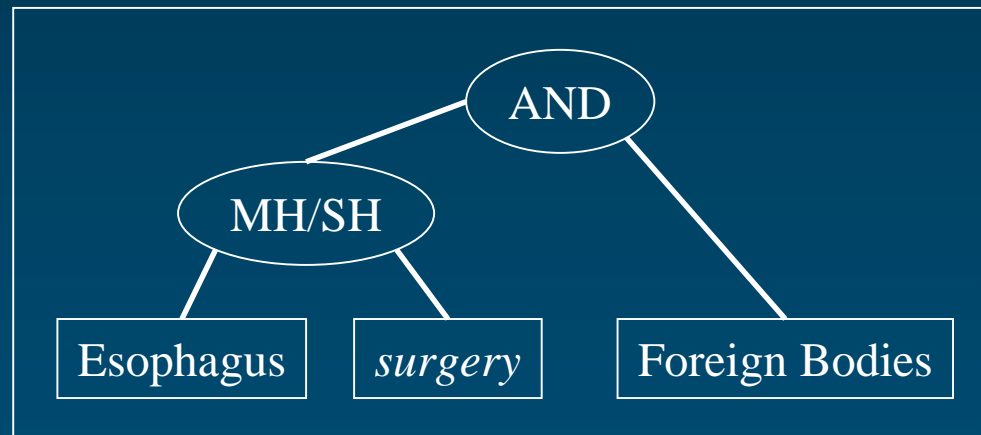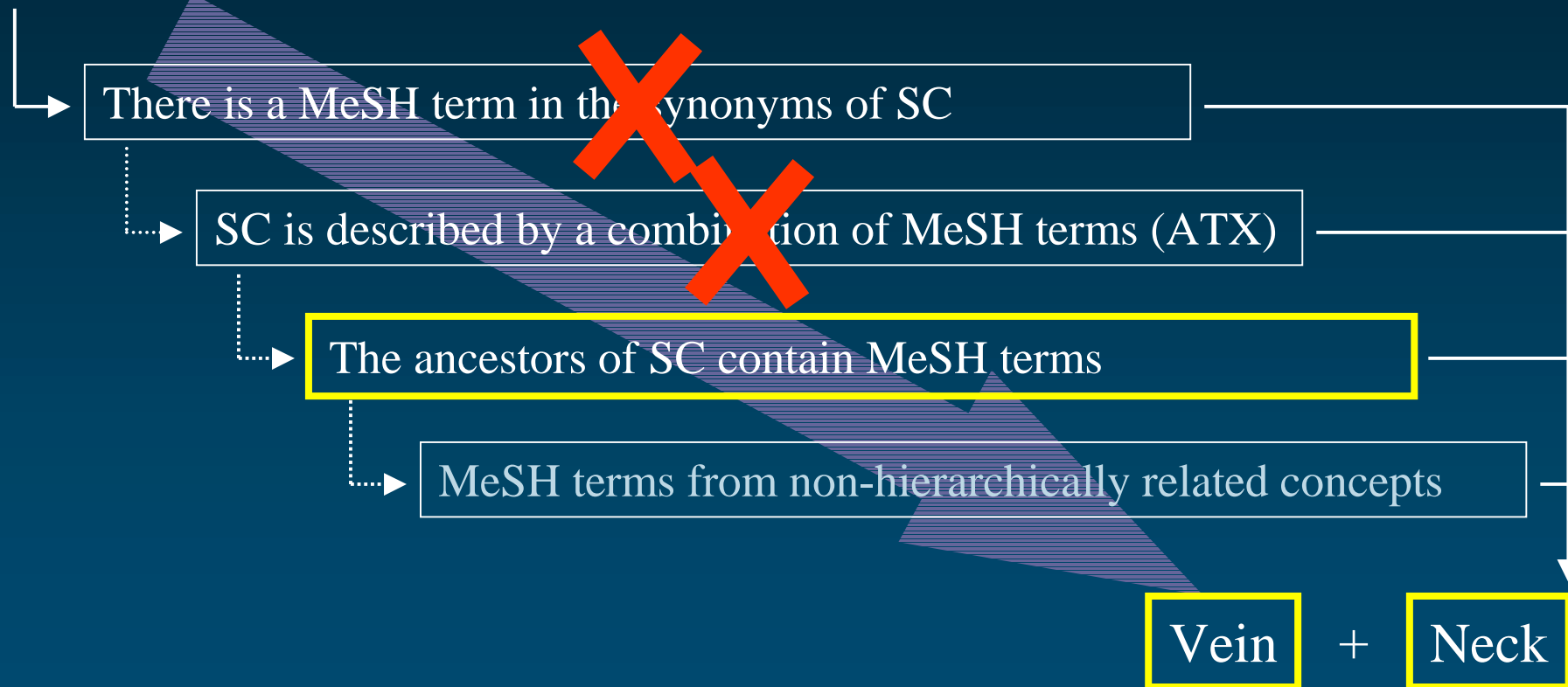
Restrict to MeSH Ancestors

- If not, let us build the graph of the ancestors of this concept
  - using parents and broader concepts (MRREL)
  - all the way to the top
  - excluding ancestors whose semantic types are not compatible with those of the source concept (MRSTY)
- From the graph, select the concepts that come from MeSH (MRSO)
- Remove those that are ancestors of another concept coming from MeSH

Restrict to MeSH Other related concepts

- If not, explore the other related concepts (MRREL) whose semantic types are compatible with those of the source concept (MRSTY)

- From those, select the concepts that come from MeSH (MRSO)

Restrict to MeSH Example

Vein of neck, NOS

There is a MeSH term in the synonyms of SC

SC is described by a combination of MeSH terms (ATX)

The ancestors of SC contain MeSH terms

MeSH terms from non-hierarchically related concepts

Vein + Neck

Restrict to MeSH Example

Head and neck, NOS

Body part, NOS

Blood Vessels

Vascular structure

Head

Veins

Neck

Systemic veins

Vein of head and neck, NOS

Vein of neck, NOS
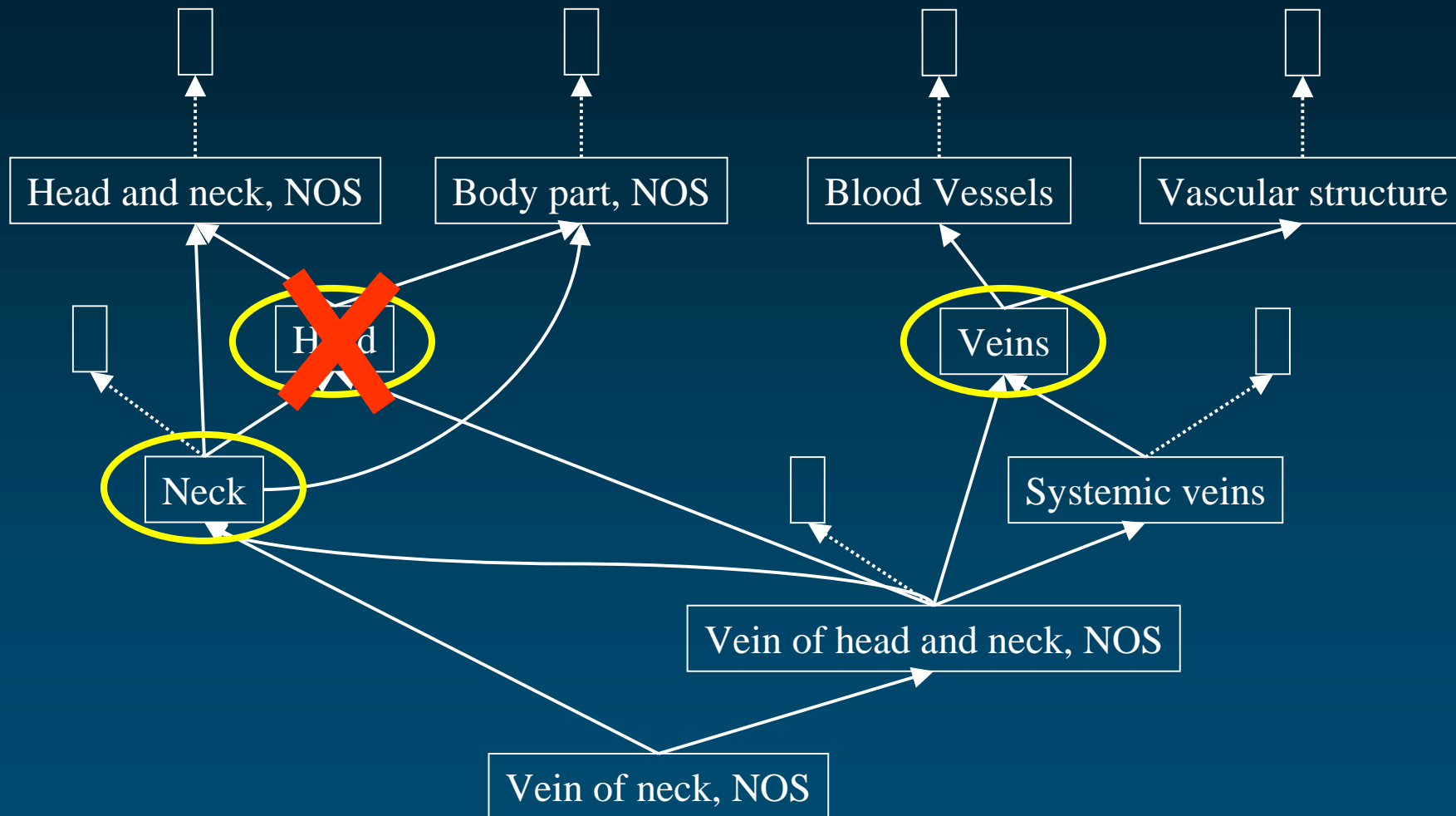
Restrict to MeSH Quantitative results

◆ 82.5% of UMLS concepts mapped to MeSH

Other related concepts

Synonymy

32%

11%

Graph of
ancestors

56%

1%

Associated
expressions
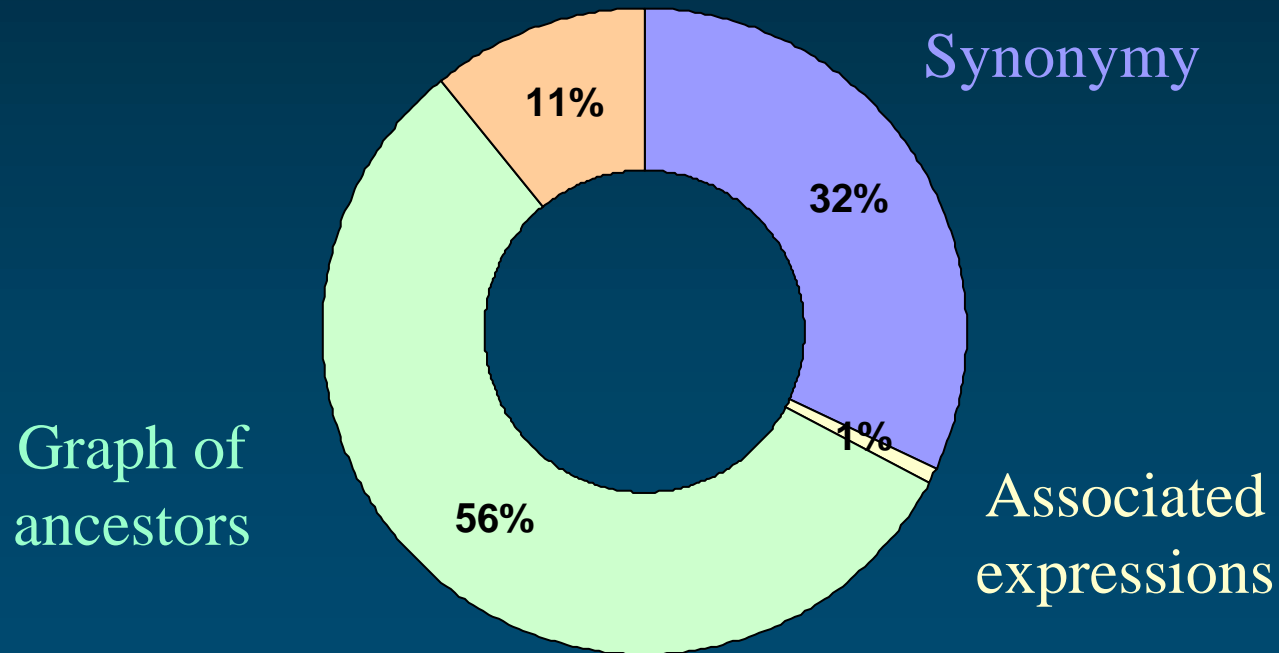
Restrict to MeSH Qualitative results

- ◆ Qualitative evaluation
  - ● 1,036 concepts extracted from 200 MEDLINE citations
  - ● manual review of every mapping or failure
- ◆ 61% Relevant
  - ● Subtotal Gastrectomy ➡ Gastrectomy
  - ● Encephalopathy, NOS ➡ Brain Diseases
- ◆ 28% More or less relevant
  - ● Vitamin A measurement ➡ Laboratory Procedure
  - ● Swelling, NOS ➡ Symptoms
- ◆ 11% Non relevant

Benefits and Limitations

Benefits

UMLS compared to individual vocabularies

- Broader scope

- Extended coverage

- Finer granularity

- Unique identifier

- Synonymous terms clustered into concepts

- Additional synonyms

- Additional hierarchical relationships

- Semantic categorization

Direct benefits

- ◆ Concept categorization
- ◆ Information retrieval
  - Synonyms
  - Cross-language features
- ◆ Information extraction
  - MetaMap
  - Normalization
- ◆ Information visualization
  - Knowledge Source Server
  - Semantic Navigator

UMLA as an enabling resource

◆ Examples

  ● Mapping across vocabularies

  ● Semantics of statistical associations

 ● Redundancy in hierarchical relations

Limitations

Limitations

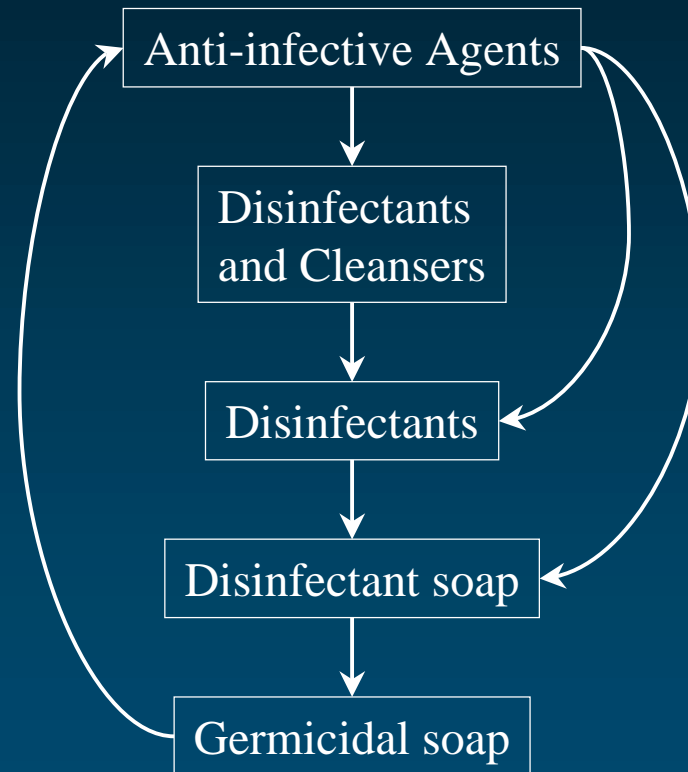- ◆ **Structural inconsistency**
  - ● Cycles in the graph of hierarchical relations
- ◆ **Semantic inconsistency**
  - ● Between Metathesaurus and Semantic Network
- ◆ **Missing relations**
  - ● Synonymy
  - ● Hierarchical relations (missing or underspecified)

Structural inconsistency From trees to graph

- Multiple tree structures combined into a graph structure
- Directed acyclic graph (DAG)
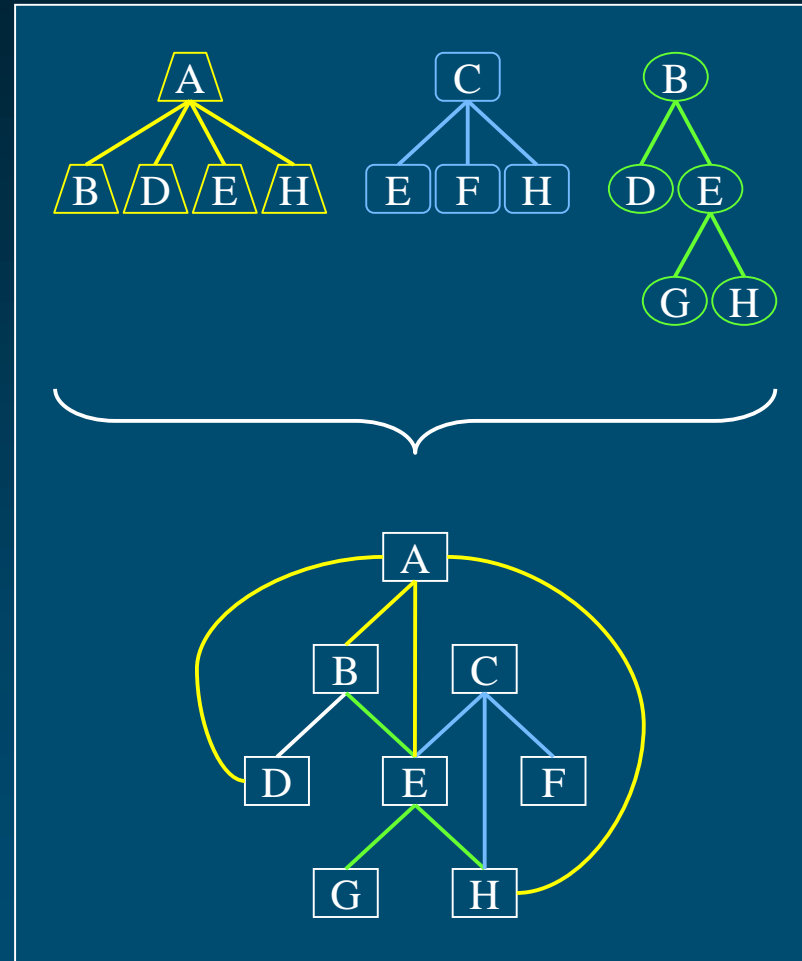
Structural inconsistency There are some cycles

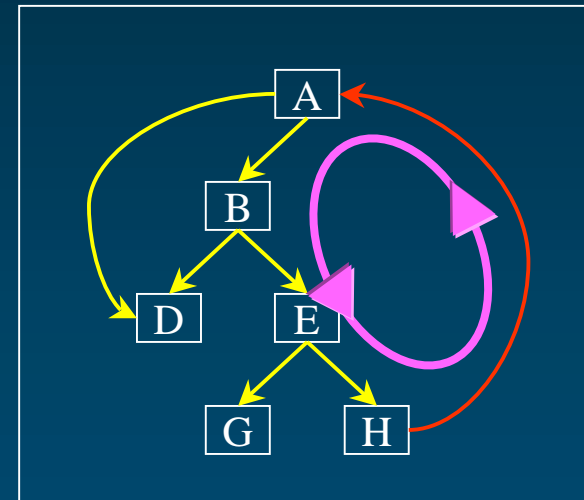# Structural inconsistency Issues

- ◆ Theoretical
  - ● Violate the antisymmetry property of partial ordering relations

- ◆ Practical
  - ● Loops in graph traversal
  - ● Impossible to perform transitive reduction

[Bodenreider, AMIA 2001]

159

Semantic inconsistency A two-level structure

Semantic Network

Fully Formed Anatomical Structure

Biologic Function

Pathologic Function

isa

location of

isa

Body Part, Organ, or Organ Component

Disease or Syndrome

isa

Adrenal Cortex

Adrenal Cortical hypofunction

location of

Metathesaurus

160
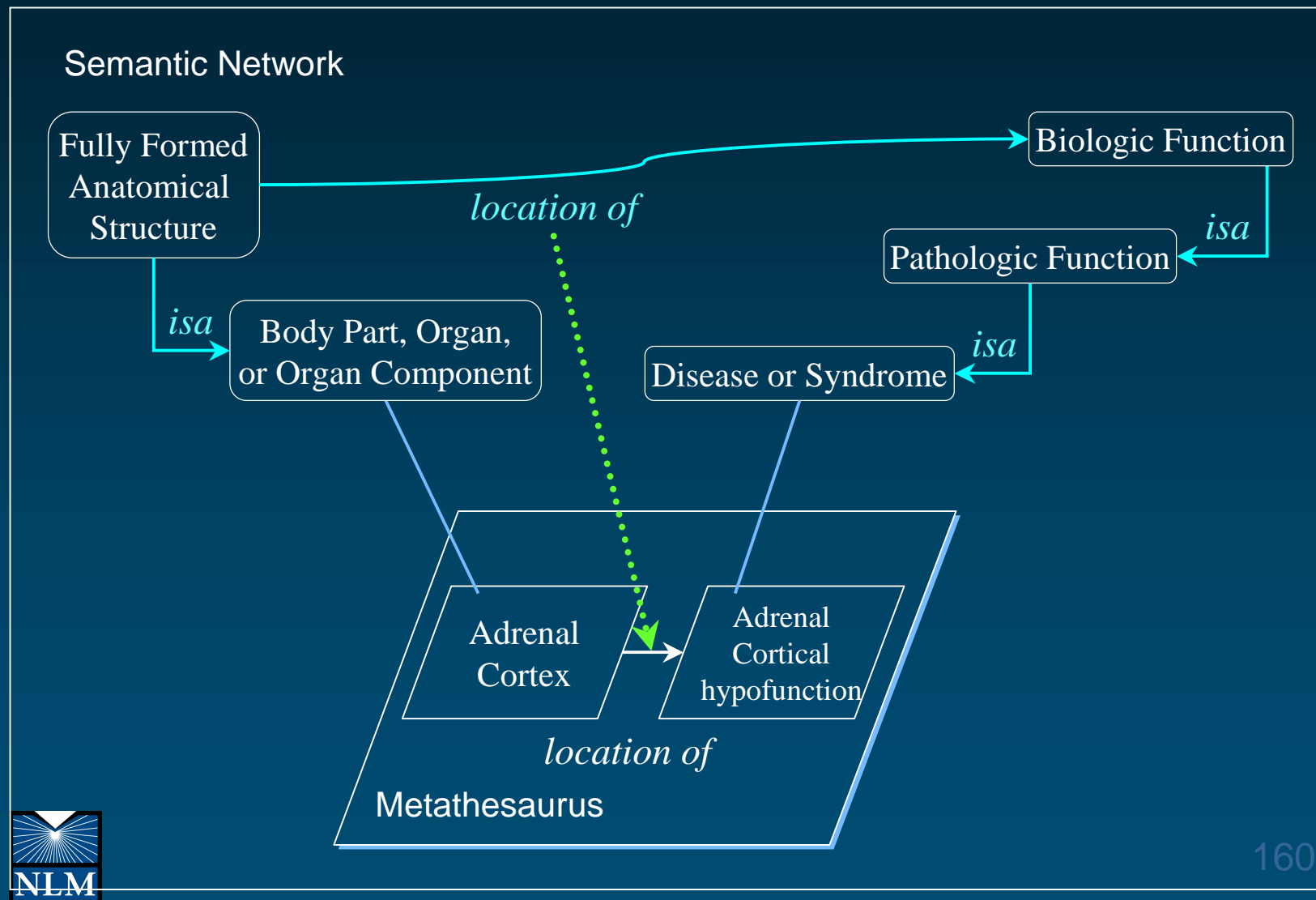
Semantic inconsistency A limited study

◆ 6894 interconcept relationships

- among the 3764 concepts in the semantic neighborhood of "Heart"

ICR and SNR not compatible

ICR = SNR or ICR descendant of SNR

ICR not specified and SNR compatible and multiple
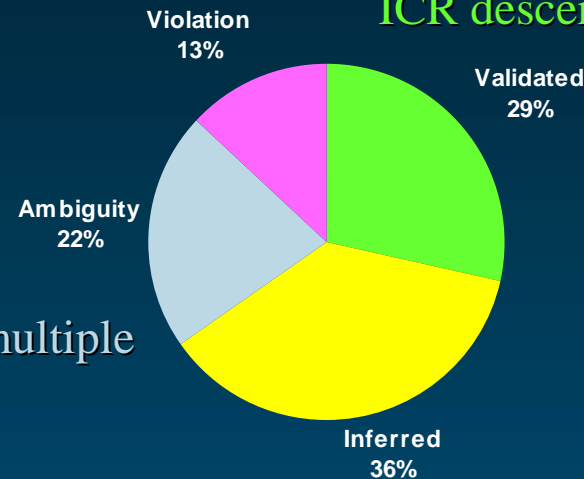
ICR not specified and SNR compatible and unique

Violation 13%

Validated 29%

Ambiguity 22%

Inferred 36%

McCray A.T, Bodenreider O. A conceptual framework for the biomedical domain. In: Green R, Bean CA, Myaeng SH, editors. *The semantics of relationships: an interdisciplinary perspective*. Boston: Kluwer Academic Publishers; 2002. p. 181-198.

NLM

Semantic inconsistency Issues

◆ The UMLS integrates what terminologies represent

◆ Hierarchies in source vocabularies
  - Often task-driven rather than based on principles
  - Usually suitable for information retrieval
  - Not necessarily suitable for reasoning

◆ No automatic correction possible
  - Wrong categorization
  - Wrong inter-concept relationship
  - [Wrong semantic network relationship]

162

Missing relations Example

diseases of the skin and subcutaneous tissues

eczema

acute eczema

infantile eczema

acute infantile eczema
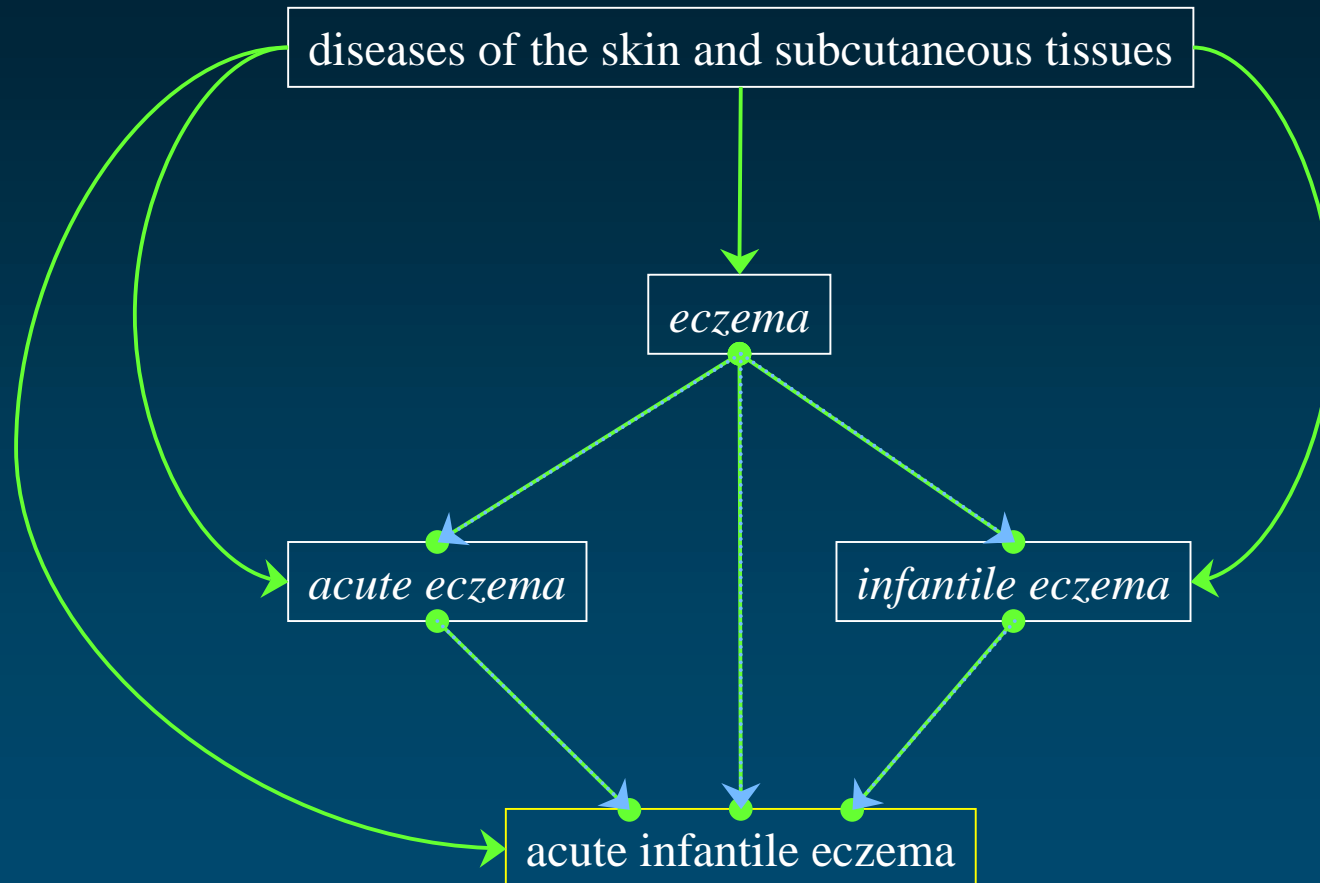
Missing relations Example
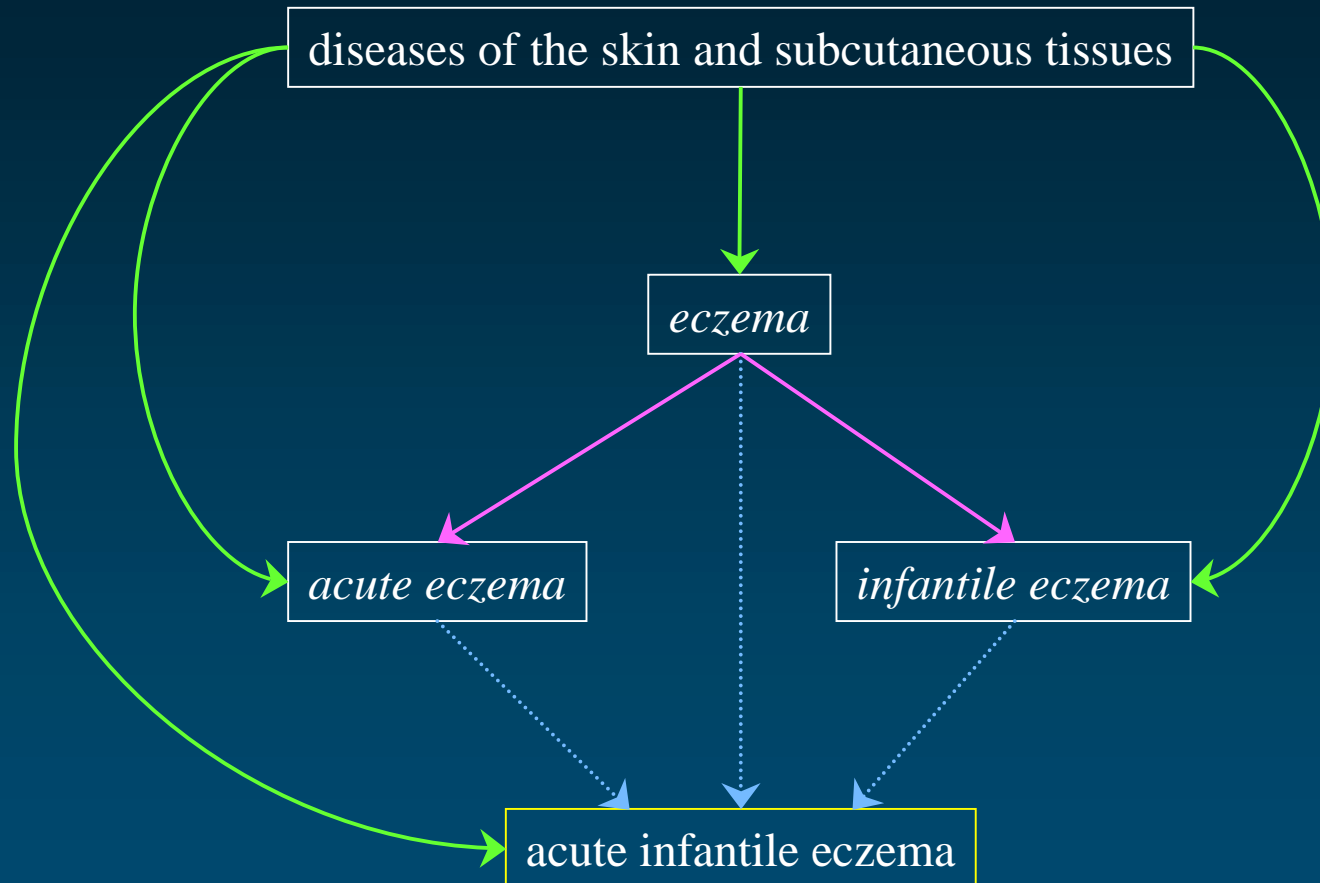
Missing relations A limited study

- 28,851 pairs of terms
  - Original SNOMED term
  - Demodified term (found in UMLS)
- Corresponding relationship in the Metathesaurus
  - Hierarchical   in 50% of the cases
  - « Sibling »   in 25% of the cases
  - Missing   in 25% of the cases

[Bodenreider & al., *TIA*, 2001]

Compensation mechanisms

◆ Examples
  ● Removing cycles from hierarchical relations
    ▪ Using redundancy (number of sources asserting the relation)
    ▪ Using terminological knowledge (e.g., NEC)
  ● Lexically-suggested hyponymic relations
    ▪ Properties of adjectival modification

More limitations

- ◆ Meaning of *isa*

- ◆ Some missing / wrong relations are hard to detect

- ◆ Some relations are present but hard to find

Meaning of *isa*

Autoimmune Diseases

is generally a

Addison's disease

Tuberculous
Addison's disease

Addison's disease
due to autoimmunity
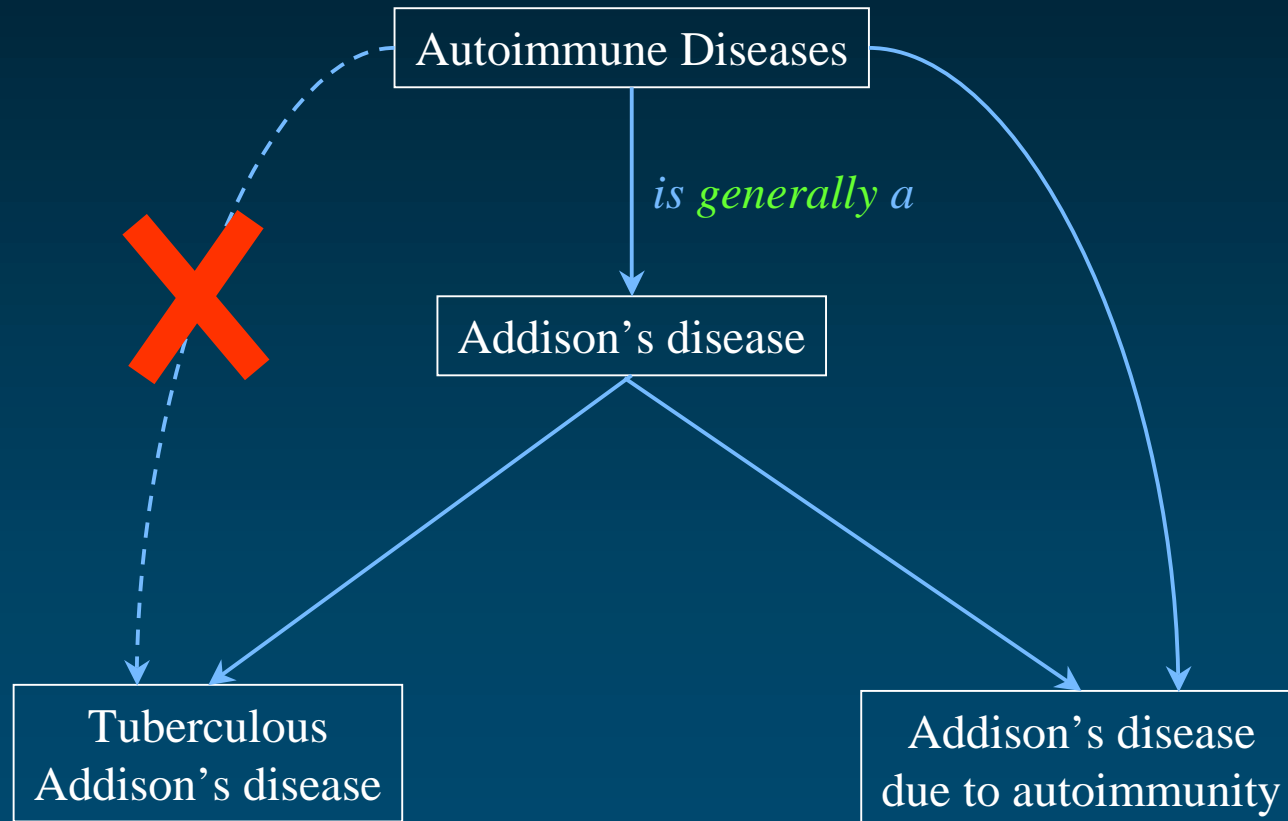
Relations Missing and difficult to detect

chronic uremia
chronic renal failure

hypertensive renal failure

chronic hypertensive uremia

NLM
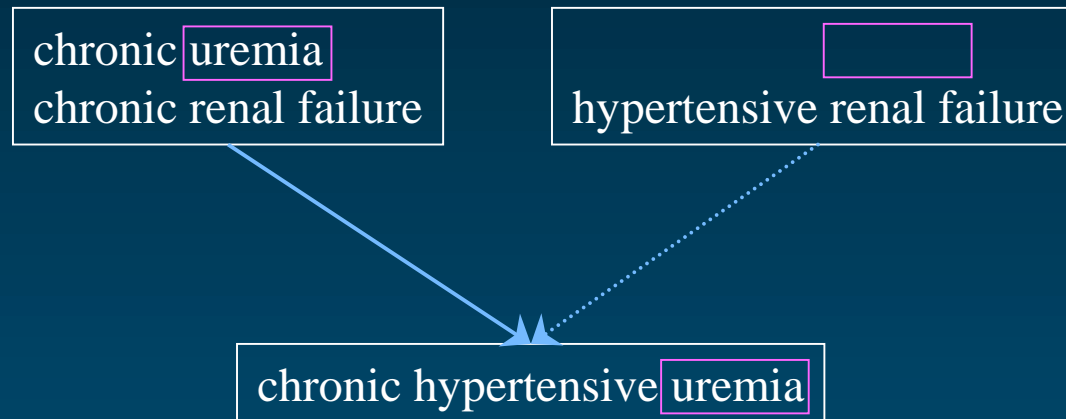
Relations Existing but difficult to find

carrier protein

↓

iron-binding protein

↓

ferritin ⟵········· iron

UMLS

Gene Ontology

has fuction

ferritin ·········⟶ iron ion transport
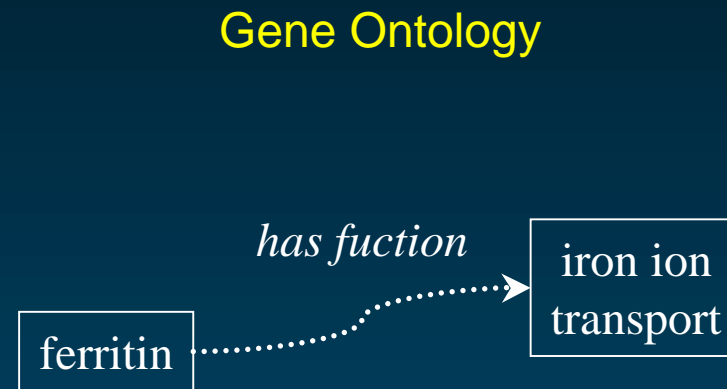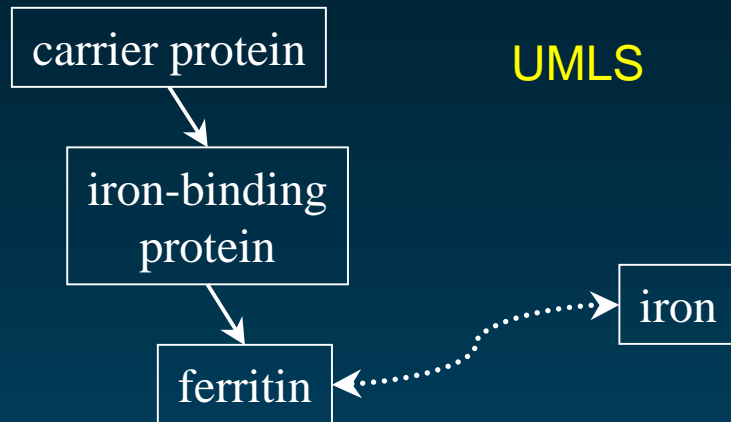
ferritin *isa* (iron transporter)

reified "transport" relationship

ferritin (*transports*) iron

"transport" relationship

How to address these limitations?

- Description logics

- Natural Language Processing
  (semantic interpretation of the terms)

- Comparing knowledge sources
  (alignment, inference)

Appendix

MRCON Concepts

(2003AA)

```
CUI        LAT  TS  LUI       STT  SUI       STR                                  LRL
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0010792|Addison Disease|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0278071|PF|S0352321|ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)|0|
C0001403|ENG|S|L0278422|PF|S0352329|ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE|0|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0368000|PF|S0496840|Melasma addisonii|3|
C0001403|ENG|S|L0368398|PF|S0506528|Primary adrenal deficiency|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
C0001403|ENG|S|L0377831|PF|S0473611|Bronzed disease|3|
C0001403|ENG|S|L0494940|PF|S0718028|Primary adrenocortical insufficiency|3|
C0001403|ENG|s|L0494937|PF|S0718027|Primary adrenocortical insuff|3|
C0001403|FIN|P|L1510041|PF|S1805950|Addisonin tauti|3|
C0001403|FRE|S|L1272481|PF|S1514427|MALADIE D'ADDISON|2|
C0001403|GER|P|L1229627|PF|S1471573|Addison-Krankheit|3|
C0001403|GER|S|L1288823|PF|S1530769|Primaere Nebennierenrindeninsuffizienz|1|
C0001403|ITA|P|L1276837|PF|S1518783|Morbo di Addison|3|
C0001403|POR|P|L0324623|PF|S0432928|DOENCA DE ADDISON|2|
C0001403|RUS|P|L0889403|PF|S1093220|ADDISONOVA BOLEZN'|3|
C0001403|SPA|P|L0342625|PF|S0450930|ENFERMEDAD DE ADDISON|3|
[…]
```
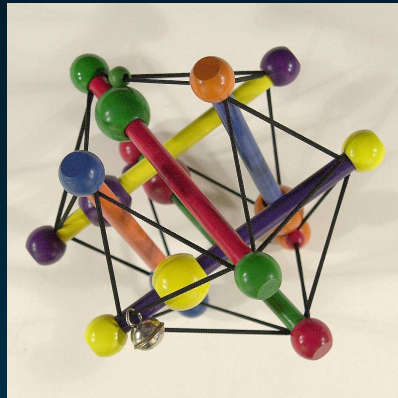
MRSO Sources

```
CUI        LUI        SUI        SAB  TTY  SCD         SRL

C0001403|L0001403|S0010792|MSH|EN|D000224|0|
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0010796|MSH|PM|D000224|0|
C0001403|L0001403|S0010796|PSY|PT|00810|3|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0220088|MSH|PM|D000224|0|
C0001403|L0001403|S0352252|CCPSS|PT|0022753|3|
C0001403|L0001403|S0352252|DXP|SY|NOCODE|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0354372|AOD|DE|0000005430|0|
C0001403|L0001403|S0354372|CSP|PT|0060-3321|0|
C0001403|L0001403|S0354372|LCH|PT|U000061|0|
C0001403|L0001403|S0354372|MDR|LT|10001130|3|
C0001403|L0001403|S0354372|RCD|PT|C1541|3|
C0001403|L0001403|S0354372|SNM|SY|D-2332|3|
C0001403|L0001403|S0365923|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0001403|S1619433|MDR|LT|10001130|3|
C0001403|L0001403|S1911394|ICPC2P|PT|T99002|3|
C0001403|L0001403|S1921523|MTHICD9|ET|255.4|0|
C0001403|L0001403|S1932462|ICPC2P|SF|T99002|3|
[…]
```

MRDEF Definitions

```
CUI      SAB      DEF

C0001403|MSH|A disease characterized by hypotension, weight loss, anorexia,
weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It
is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the
adrenal glands that results in deficiency of aldosterone and cortisol. In the
absence of replacement therapy, it is usually fatal.|
[…]
```

MRSTY Semantic Types

(2003AA)

```
CUI      TUI  STY
C0001400|T040|Organism Function|
C0001403|T047|Disease or Syndrome|
C0001406|T083|Geographic Area|
C0001407|T114|Nucleic Acid, Nucleoside, or Nucleotide|
C0001407|T123|Biologically Active Substance|
[…]
```

MRATX Associated Expressions (2003AA)

```
CUI        SAB REL      ATX
```

Closed fracture of malar and maxillary bones, NOS
`C0009045|MSH|RB|<Zygomatic Fractures> OR <Maxillary Fractures>|`

Unilateral congenital dislocation of hip
`C0009702|MSH|RB|<Hip Dislocation, Congenital> AND <Femur Head>/<abnormalities>|`

Suture of bladder
`C0010700|MSH|RB|<Bladder>/<surgery>|`

Corneal abrasion
`C0010032|MSH|RO|<Cornea>/<injuries>|`

CORRECTIVE LENS PROBLEM
`C0010099|MSH|RO|<Contact Lenses>/<adverse effects>|`

Chronic cough
`C0010201|MSH|SY|<Cough> AND <Chronic Disease>|`

Cyst and pseudocyst of pancreas
`C0010623|MSH|SY|<Pancreatic Cyst> OR <Pancreatic Pseudocyst>|`

Cystitis
`C0010692|LCH|RU|<Bladder>/<Inflammation>|`
`[…]`

MRCXT Contexts

```
CUI        SUI      SAB   SCD       CXN CXL RNK   CXS                          CUI2       HCD REL XC
C0001403|S0469271|SNMI|DB-70620|1|ANC|1|SNOMED International|C1140118||||
C0001403|S0469271|SNMI|DB-70620|1|ANC|2|DISEASES/DIAGNOSES|C0338067||||
C0001403|S0469271|SNMI|DB-70620|1|ANC|3|DISEASES OF THE END. SYSTEM|C0014130||||
C0001403|S0469271|SNMI|DB-70620|1|ANC|4|DISEASES OF THE ADRENAL GLANDS|C0001621||||
C0001403|S0469271|SNMI|DB-70620|1|CCP||Addison's disease, NOS|C0001403|DB-70620|||

(* = C0001403|S0718028|ICD10)
*|E27.1|1|ANC|1|ICD…, Tenth Revision (ICD-10)|C1140143||||
*|E27.1|1|ANC|2|Endocrine, nutritional and metabolic diseases|C0694452|E00-E90.9|||
*|E27.1|1|ANC|3|Disorders of other endocrine glands|C0178257|E20-E35.9|||
*|E27.1|1|ANC|4|Other disorders of adrenal gland|C0494313|E27|||
*|E27.1|1|CCP||Primary adrenocortical insufficiency|C0001403|E27.1|||

(* = C0001403|S0010794|MSH)
*|D000224|1|ANC|1|MeSH|C1135584||||
*|D000224|1|ANC|2|MeSH Descriptors|C1135587||||
*|D000224|1|ANC|3|Index Medicus Descriptor|C1135589||||
*|D000224|1|ANC|4|Diseases (MeSH Category)|C0012674|C|||
*|D000224|1|ANC|5|Endocrine Diseases|C0014130|C19|||
*|D000224|1|ANC|6|Adrenal Gland Diseases|C0001621|C19.53|||
*|D000224|1|ANC|7|Adrenal Gland Hypofunction|C0001623|C19.53.264|||
*|D000224|1|CCP||Addison's Disease|C0001403|C19.53.264.263|||
*|D000224|1|SIB||Adrenoleukodystrophy|C0001661|C19.53.264.270|||
*|D000224|1|SIB||Hypoaldosteronism|C0020595|C19.53.264.480|||
```

MRSAT Simple concept attributes (2003AA)

```
CUI       LUI       SUI       SCD       ATN SAB     ATV
C0001403|L0001403|S0010792|D000224|DID|MSH|D000224|
C0001403|L0001403|S0010792|D000224|EV|MSH|ADDISON DIS|
C0001403|L0001403|S0010792|D000224|MUI|MSH|M0000346|
C0001403|L0001403|S0010792|D000224|TH|MSH|UNK (19XX)|
C0001403|L0001403|S0010794|D000224|AN|MSH|an autoimmune dis with adrenal hypofunction|
C0001403|L0001403|S0010794|D000224|AQL|MSH|BL CF CI CL CN CO DH DI DT EC EH EM EN …|
C0001403|L0001403|S0010794|D000224|DC|MSH|1|
C0001403|L0001403|S0010794|D000224|DID|MSH|D000224|
C0001403|L0001403|S0010794|D000224|EV|MSH|ADDISON DIS|
C0001403|L0001403|S0010794|D000224|MDA|MSH|19990101|
C0001403|L0001403|S0010794|D000224|MED1963|NLM-MED|*2|
C0001403|L0001403|S0010794|D000224|MED1963|NLM-MED|2|
[…]
C0001403|L0001403|S0010794|D000224|MED2002|NLM-MED|*19|
C0001403|L0001403|S0010794|D000224|MED2002|NLM-MED|23|
[…]
C0001403|L0001403|S0010794|D000224|MN|MSH|C19.53.264.263|
C0001403|L0001403|S0010794|D000224|MN|MSH|C20.111.163|
[…]
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
[…]
C0001403||||DA|MTH|19900930|
C0001403||||MR|MTH|20021026|
C0001403||||ST|MTH|R|
```

NLM

MRLO Locators

```
CUI        ISN FR  UN         SUI    SNA SOUI
C0001403|DXP|||S0352252|||
C0001403|DXP|||S0352329|||
C0001403|MBD|182|*CITATIONS|S0010794|||
C0001403|MED|179|*CITATIONS|S0010794|||
```

MRRANK Name Ranking

```
RANK SAB TTY SUPRES
0401|MTH|PN|N|
0400|MTH|MM|N|
0399|MSH|MH|N|
0398|MSH|TQ|N|
0397|MSH|EP|N|
0396|MSH|EN|N|
0395|MSH|XQ|N|
0394|MSH|NM|N|
0393|RXNORM|SCD|N|
0392|RXNORM|SCDC|N|
0391|DSM4|PT|N|
0390|DSM3R|PT|N|
0389|SNMI|PT|N|
0388|SNMI|PX|Y|
0387|SNMI|HT|N|
0386|SNMI|HX|Y|
0385|VANDF|CD|N|
0384|VANDF|HT|N|
0383|VANDF|IN|N|
0382|MDDB|CD|N|
0381|MMX|CD|N|
0380|MMX|IN|N|
0379|RCDSA|PT|N|
[…]
```

MRREL Inter-concept Relationships
(2003AA)

```
CUI1      REL CUI2      RELA       SAB   SL      MG
C0001403|AQ|C0348026||MSH|MSH||
C0001403|CHD|C0342477||RCD|RCD||
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0001621||PSY|PSY||
C0001403|PAR|C0001621||SNMI|SNMI||
C0001403|PAR|C0001623||MSH|MSH||
C0001403|PAR|C0935495|has_member|PSY|PSY||
C0001403|RB|C0001621||PSY|PSY||
C0001403|RB|C0001623||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RB|C0004364||MTH|MTH||
C0001403|RL|C0405580|mapped_from|SNMI|SNMI||
C0001403|RN|C0518933||MTH|MTH||
C0001403|RN|C0518934||MTH|MTH||
C0001403|RO|C0152889|associated_with|SNMI|SNMI||
C0001403|RO|C0546992||MTH|MTH||
C0001403|RQ|C0020615|clinically_associated_with|CCPSS|CCPSS||
C0001403|RQ|C0151467|clinically_similar|RAM|RAM||
C0001403|RQ|C0300942|classifies|MDR|MDR||
C0001403|RQ|C0405580|mapped_from|CST|CST||
C0001403|RQ|C0405580|mapped_to|HLREL|HLREL||
C0001403|RQ|C0740740|inverse_isa|CCPSS|CCPSS||
C0001403|SIB|C0001206||MDR|MDR||
[…]
```

MRCOC Co-occurrences

```
CUI1      CUI2      SOC COT COF COA
C0001403|C0000727|MED|L|1|CO=1,DI=1,ME=1|
C0001403|C0000737|MBD|L|1|CO=1,DI=1|
C0001403|C0000833|MED|L|2|MI=2,DT=1,RA=1|
C0001403|C0001175|MBD|L|1|CO=1|
C0001403|C0001418|MED|L|1|ET=1|
C0001403|C0001430|MBD|L|1|BL=1,CO=1|
C0001403|C0001551|MED|L|3|DT=3|
C0001403|C0001613|MBD|L|6|ET=2,IM=2,CL=1,CN=1,DI=1,PA=1,PP=1|
C0001403|C0001613|MED|L|6|IM=4,PP=3,CO=2,BL=1,DI=1,TH=1|
C0001403|C0001614|MBD|L|1|BL=1,CI=1|
C0001403|C0001617|MBD|L|1|BL=1|
C0001403|C0001618|MBD|L|2|BL=2,CO=1,ET=1|
C0001403|C0001618|MED|L|1|CO=1,PA=1|
[…]
C0018099|C0151373|AIR|KP|||
C0018099|C0151407|AIR|KP|||
C0018099|C0151463|CCPSS|PP|1||
C0018099|C0205082|CCPSS|MP|1||
C0018099|C0205090|CCPSS|MP|8||
C0018099|C0205091|CCPSS|MP|2||
C0018099|C0221598|AIR|KP|||
[…]
```

MRCON Suppressible synonyms   (2003AA)

```
CUI        LAT TS  LUI       STT   SUI     STR                              LRL
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0010792|Addison Disease|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0278071|PF|S0352321|ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)|0|
C0001403|ENG|S|L0278422|PF|S0352329|ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE|0|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0368000|PF|S0496840|Melasma addisonii|3|
C0001403|ENG|S|L0368398|PF|S0506528|Primary adrenal deficiency|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
C0001403|ENG|S|L0377831|PF|S0473611|Bronzed disease|3|
C0001403|ENG|S|L0494940|PF|S0718028|Primary adrenocortical insufficiency|3|
C0001403|ENG|S|L0494937|PF|S0718027|Primary adrenocortical insuff|3|
C0001403|FIN|P|L1510041|PF|S1805950|Addisonin tauti|3|
C0001403|FRE|S|L1272481|PF|S1514427|MALADIE D'ADDISON|2|
C0001403|GER|P|L1229627|PF|S1471573|Addison-Krankheit|3|
C0001403|GER|S|L1288823|PF|S1530769|Primaere Nebennierenrindeninsuffizienz|1|
C0001403|ITA|P|L1276837|PF|S1518783|Morbo di Addison|3|
C0001403|POR|P|L0324623|PF|S0432928|DOENCA DE ADDISON|2|
C0001403|RUS|P|L0889403|PF|S1093220|ADDISONOVA BOLEZN'|3|
C0001403|SPA|P|L0342625|PF|S0450930|ENFERMEDAD DE ADDISON|3|
[…]
```

MRCUI Concept history

```
CUI1      VER    CREL  CUI2   MAPIN
C0241779|1996AA|SY|C0001403|Y|
C0271735|1996AA|SY|C0001403|Y|
[…]
```

MRSAB Source information (2003AA)

```
VCUI      RCUI       VSAB      RSAB SON SF SVER MSTART MEND IMETA RMETA SLC SCC SRL TFR
C1140103|C1140104|INS2002|INS|French translation of the Medical Subject Headings,
2002|MSH|2002|2002_04_11||2002AB||Dr. Annie Advocat; e-mail: advocat@inserm-
dicdoc.u-strasbg.fr|Dr. Annie Advocat; e-mail: advocat@inserm-dicdoc.u-
strasbg.fr|3|30883|20692||MH,SY||FRE|ISO646-US|Y|Y|

C1140132|C1140133|BRMP2002|BRMP|Portuguese translation of the Medical Subject
Headings, 2002|MSH|2002|2001_12_04||2002AA||Elenice de Castro; e-mail:
elenice@brm.bireme.br|Elenice de Castro; e-mail:
elenice@brm.bireme.br|3|41853|27195||EP,MH,SY||POR|ISO646-US|Y|Y|

C1140297|C1140298|DUT2001|DUT|Dutch Translation of the Medical Subject Headings,
2001|MSH|2001|2001_12_04||2002AB||A.J.P.M.Overbeke, overbeke@ntvg.nl, * 20 662
0150|A.J.P.M.Overbeke, overbeke@ntvg.nl, * 20 662
0150|3|35705|17733||EP,MH,SY||DUT|ISO646-US|Y|Y|

C1142630|C1135584|MSH2003_2002_10_24|MSH|Medical Subject Headings,
2002_10_24|MSH|2003_2002_10_24|2002_11_05||2003AA||Stuart Nelson, M.D., Head, MeSH
Section; e-mail: nelson@nlm.nih.gov|Stuart Nelson, M.D., Head, MeSH Section; e-
mail: nelson@nlm.nih.gov|0|516015|231458|FULL-
MULTIPLE|CE,EN,EP,HS,HT,MH,N1,NM,PM,TQ,XQ|AN,AQL,CX,DC,DID,DQ,DS,DX,EC,EV,FR,FX,HM
,HN,II,LT,MDA,MMR,MN,MUI,OL,PA,PI,PM,QA,QE,QS,RN,RR,SOS,SRC,TH|ENG|ISO646-US|Y|Y|
```

SRDEF Basic information

(2003AA)

```
RT   TUI   STY/RL   STN/RTN   DEF      EX      UN      NH      ABR      RIN
STY|T001|Organism|A1.1|Generally, a living individual, including all plants and
animals.|Homozygote; Radiation Chimera; Sporocyst|||||
STY|T002|Plant|A1.1.1|An organism having cellulose cell walls, growing by
synthesis of inorganic substances, generally distinguished by the presence of
chlorophyll, and lacking the power of locomotion.  Plant parts are included here
as well.|Pollen; Potatoes; Vegetables|||||
STY|T003|Alga|A1.1.1.1|A chiefly aquatic plant that contains chlorophyll, but does
not form embryos during development and lacks vascular tissue.|Chlorella;
Laminaria; Seaweed|||||
STY|T004|Fungus|A1.1.2|A eukaryotic organism characterized by the absence of
chlorophyll and the presence of a rigid cell wall.  Included here are both slime
molds and true fungi such as yeasts, molds, mildews, and mushrooms.|Aspergillus
clavatus; Blastomyces; Helminthosporium; Neurospora|||||
[…]
RL|T132|physically_related_to|R1|Related by virtue of some physical attribute or
characteristic.||||PR|physically_related_to|
RL|T133|part_of|R1.1|Composes, with one or more other physical units, some larger
whole.  This includes component of, division of, portion of, fragment of, section
of, and layer of.||||PT|has_part|
[…]
RL|T186|isa|H|The basic hierarchical link in the Network.  If one item "isa"
another item then the first item is more specific in meaning than the second
item.|||||IS|inverse_isa|
[…]
```

SRSTR Structure

```
STY/RL                    RL        STY/RL                                              LS
Biologic Function|affects|Organism|D|
Biologic Function|isa|Natural Phenomenon or Process|D|
Biologic Function|process_of|Organism|D|
Biologic Function|produces|Biologically Active Substance|D|
Biologic Function|produces|Body Substance|D|
[…]
Disease or Syndrome|conceptually_related_to|Experimental Model of Disease|DNI|
Disease or Syndrome|isa|Pathologic Function|D|
Disease or Syndrome|produces|Tissue|D|
[…]
Medical Device|isa|Manufactured Object|D|
Medical Device|prevents|Injury or Poisoning|D|
Medical Device|prevents|Pathologic Function|D|
Medical Device|treats|Anatomical Abnormality|D|
Medical Device|treats|Injury or Poisoning|D|
Medical Device|treats|Pathologic Function|D|
Medical Device|treats|Sign or Symptom|D|
[…]
Mental Process|process_of|Plant|B|  blocks  Biologic Function|process_of|Organism|D|
[…]
part_of|isa|physically_related_to|D|
[…]
```

SRSTRE2 Structure (expanded) (2003AA)

```
STY                 RL          STY
Disease or Syndrome|isa|Pathologic Function|          Pathologic Function |isa|Biologic Function|
Disease or Syndrome|isa|Biologic Function|            Biologic Function|isa|Natural Phen. or Process|
Disease or Syndrome|isa|Natural Phen. or Pr.|         Natural Phen. or Process|isa|Phen. or Process|
Disease or Syndrome|isa|Phenomenon or Process|        Phenomenon or Process|isa|Event|
Disease or Syndrome|isa|Event|
Disease or Syndrome|affects|Alga|
Disease or Syndrome|affects|Amphibian|
Disease or Syndrome|affects|Animal|
Disease or Syndrome|affects|Archaeon|
Disease or Syndrome|affects|Bacterium|                from  Biologic Function|affects|Organism|D|
Disease or Syndrome|affects|Biologic Function|
Disease or Syndrome|affects|Bird|
Disease or Syndrome|affects|Cell Function|
Disease or Syndrome|affects|Cell or Molecular Dysfunction|
[…]
```

Bibliography

References: UMLS home page

◆ UMLS home page

> http:// www.nlm.nih.gov/research/umls/

◆ UMLS documentation

- "Green Book"
- online documentation

> http://www.nlm.nih.gov/research/umls/UMLSDOC.HTML

◆ UMLS Information web site

> http://umlsinfo.nlm.nih.gov/

References

- ◆ UMLS as a research project
  - Lindberg, D. A., Humphreys, B. L., & McCray, A. T. (1993). The Unified Medical Language System. *Methods Inf Med, 32*(4), 281-91.
  - Humphreys, B. L., Lindberg, D. A., Schoolman, H. M., & Barnett, G. O. (1998). The Unified Medical Language System: an informatics research collaboration. *J Am Med Inform Assoc, 5*(1), 1-11.

NLM

References

- Technical papers
  - McCray, A. T., & Nelson, S. J. (1995). The representation of meaning in the UMLS. *Methods Inf Med, 34*(1-2), 193-201.
  - Campbell, K. E., Oliver, D. E., Spackman, K. A., & Shortliffe, E. H. (1998). Representing thoughts, words, and things in the UMLS. *J Am Med Inform Assoc, 5*(5), 421-31.
- Comprehensive bibliography 1986-96

  http://www.nlm.nih.gov/pubs/cbm/umlscbm.html